

STIC Biotech/ChemLib

76549

Fr m: Rao, Manjunath N.
Sent: Thursday, September 26, 2002 11:05 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/870,113

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10C 01
Phone: 306-5681

CRIFE

Date: 9-20-02

Please search the following as soon as possible for application with serial number **09/870,113**

SEQ ID NO:3, 5,7,9,11, against all commercial nucleic acid databases including issued patents database and pending application database and provide a **print of all results**.

If y u have any questions please call me at the above phone number.

Thanks

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Manjunath N. Rao, Ph.D.
Biotechnology Patent Examiner
Art Unit 1652, Room 10A11
Crystal Mall 1, USPTO.

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 9/26/02
Date Completed: 9/29/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH: 5
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: PS
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 03:51:19 ; Search time 5287.54 Seconds
(without alignments)
2303.386 Million cell updates/sec

Title: us-09-870-113-3
Perfect score: 582
Sequence: 1 atggagttgagggcgggg.....tgaaacctgcgaaggtga 582

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	577	99.1	1448	9	AF327402
2	577	99.1	1448	17	HSA303077
3	520.4	89.4	1889	9	AF327403
4	520.4	89.4	1889	17	HSA303078
5	454.4	78.1	1322	6	AX061229
c	291	50.0	123160	9	AL353719
7	273	46.9	401	6	AX071440
8	251	43.1	221062	2	AC096351
c	210.6	36.2	221062	2	AC096351
10	207.2	35.6	4112	10	AF288621
11	198.8	34.2	1429	9	AF223466
12	177.8	30.5	5869	10	AF361699
13	174.2	29.9	3415	9	AY032628
14	172	29.6	858	9	BC015013
15	125.2	21.5	163205	2	AC099368
16	116.2	20.0	191702	2	AC108878
17	113.8	19.6	108765	2	AF216674
c	112.2	19.3	174445	9	AC051642
c	112.2	19.3	182504	2	AC022597
19	105.2	18.1	483	6	AX306258
20	82.4	14.2	34305	3	CEW02B12
c	72.4	12.4	3037	3	AY060268
22	71.6	12.3	146743	2	AF003518
c	70.8	12.2	2878	3	AF217402
c	68	11.7	615	11	CNS06FGL
26	68	11.7	71053	2	AC102327
c	67.2	11.5	77191	2	AC023381
28	67.2	11.5	207420	2	AC078884
29	67	11.5	121146	2	AC108558
30	66.6	11.4	38723	2	AC095142
31	66.6	11.4	59632	2	AC102213
32	66.6	11.4	94914	2	AC105592
33	66.2	11.4	158167	2	AC034105
34	66	11.3	65928	2	AC012570
35	65.8	11.3	79355	2	AC021280
c	65.6	11.3	46181	2	AC105855
37	65.6	11.3	115715	2	AC079411
38	65.6	11.3	140469	2	AC025066
39	65.6	11.3	148407	2	AC106354
c	65.2	11.2	1560	14	HS21E
41	65.2	11.2	62649	2	AC022552
42	65.2	11.2	146744	2	AC098461
43	65.2	11.2	174096	2	AC097177
44	65	11.2	93663	2	AC107039
45	64.8	11.1	60984	2	AC102552

ALIGNMENTS

RESULT 1

AF327402

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AF327402 1448 bp mRNA linear PRI 02-MAY-2001
Homo sapiens putative mitochondrial solute carrier splice variant
mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.
AF327402
AF327402.1 GI:13926046
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1448)
Li.F.Y., Nikali.K., Grogan.J., Leibiger.I., Leibiger.B.,
Schweyen.R., Larsson.C. and Suomalainen.A.
Characterization of a novel human putative mitochondrial
transporter homologous to the yeast mitochondrial RNA splicing
proteins 3 and 4
FEBS Lett. 494 (1-2), 79-84 (2001)
21195335
11297739
2 (bases 1 to 1448)

Best Local Similarity 100.0%; Pred. No. 3.1e-83;		Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	atgaggtgagggcgggggtgctgcggtgtgctgagggggggccggcgagggcccggg	60
Db	30	ATGAGTGTGAGGGCGGGGTGTGGCGGTGTGCGGGGGGGCGGCGAGGCCCGGG	89
QY	61	cgaagcccccgggaagtcgcgcctgtgacgggtgctgcagcggggcgtgagccgagg	120
Db	90	CGAGACCCGGGNGTGGCGCTGCTGGACGGGTGCTGACCGGGCGTGGGCGGGGG	149
QY	121	gccggcgcggggagggccgggggacctgagggcccccgtacgacaagatccgggactccggc	180
Db	150	GCGGCGCGGGAGGGCGGGGGCTGCAAGGCCCGCGGTACGACAAAGATCCGGACTCCGC	209
QY	181	ccggactcacgagcgctgcgcgtcgagccacatgtaccacgacacatggtggaaggcc	240
Db	210	CCGGACTACGAGCGCTGCCGGCTGGAGCACTGTACCACGCACATGGTGGCAGGGGCC	269
QY	241	gtggcagggatcctgggagcactgcgtgatgcccccactgcactgcgtcaagacccgatg	300
Db	270	GTGGCAGGGATCTTGGAGCACTGCGTGTATGCCCATCGACTCGCTCAAGACCCGGATG	329
QY	301	caagagtacagcctgaccagatgccggtatgcgaatgtgtggaggccctctggagg	360
Db	330	CAGAGTCTACAGCTGACCCAGTGCCTGTGCAATGTGTGGAGGGCCCTCTGGAGG	389
QY	361	attataagaacgagggccctatggaggcccatgaggggctgaacgtcacagcaacggc	420
Db	390	ATTATAAGAACGGAGGGCCCTATGGAGGCCCATGAGGGGGCTGAACGTCACAGCAACAGGC	449
QY	421	gcaggggcctgccacgcccctttatttgcctgtctacgaaagttaaaaaagacattgagt	480
Db	450	GCAGGGCTCCCCACGCCCTTTATTTTTCGTGTACGAAAGTTAAAAAAGACATTCAGT	509
QY	481	gatgaatccaccctggggcgaatagccatattgccaatgtgcggccgagtgatggca	540
Db	510	GATGTAATCCACCCTGGGGCAATAGCCATATTGCCAATGGTGGCGCCGGGTGTGGCA	569
QY	541	acattactcatgatgcagccatgaacccctgcggaag	577
Db	570	ACATTACTTCATGATCAGCATGAAACCCCTGCGGAAG	606

REFERENCE
AUTHORS
Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimege,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisege,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogoh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleceyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GEZD
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-tp estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
32851: contig of 32851 bp in length
32852: gap of unknown length
32951: contig of 16033 bp in length
32952: gap of unknown length
49885: gap of unknown length
49886: contig of 14731 bp in length
63815: gap of unknown length
63916: contig of 12090 bp in length
63916: gap of unknown length
76105: contig of 12736 bp in length
88841: gap of unknown length
88842: contig of 11650 bp in length
100591: gap of unknown length
100592: contig of 11011 bp in length
111702: gap of unknown length
111802: contig of 10862 bp in length
111803: gap of unknown length
122665: contig of 6947 bp in length
122765: gap of unknown length
129712: contig of 7436 bp in length
129712: gap of unknown length
137247: contig of 7329 bp in length
137248: gap of unknown length
144676: contig of 7329 bp in length
144677: gap of unknown length
152124: contig of 7348 bp in length
152224: gap of unknown length
158878: contig of 6654 bp in length
158879: gap of unknown length
163928: contig of 4950 bp in length
163929: gap of unknown length
164028: contig of 5325 bp in length
169353: gap of unknown length
169354: contig of 6137 bp in length
169454: gap of unknown length
175590: contig of 4785 bp in length
175591: gap of unknown length
180475: contig of 4082 bp in length
180476: gap of unknown length
184576: contig of 5575 bp in length
184577: gap of unknown length
184578: contig of 3702 bp in length
190333: gap of unknown length
194135: contig of 2358 bp in length
194235: gap of unknown length
196592: contig of 3838 bp in length
196593: gap of unknown length
200530: contig of 2179 bp in length
200531: gap of unknown length
202810: contig of 2125 bp in length
202810: gap of unknown length
205034: contig of 2042 bp in length
205035: gap of unknown length
207176: contig of 1569 bp in length
207177: gap of unknown length
208845: contig of 2326 bp in length
208846: gap of unknown length
211271: contig of 1931 bp in length
211272: gap of unknown length
213302: contig of 2082 bp in length
213303: gap of unknown length
215485: contig of 1385 bp in length
215485: gap of unknown length
216959: contig of 1289 bp in length
216970: gap of unknown length
218358: contig of 1428 bp in length
218359: gap of unknown length
219886: contig of 1076 bp in length
219887: gap of unknown length
219987: contig of 1076 bp in length.
Location/Qualifiers
1. .221062
/organism="Rattus norvegicus"

FEATURES
source


```

info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 15 Row: h Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7706149.

FEATURES
    Location/Qualifiers
1..858
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="MGC:8764 IMAGE:3910567"
     /tissue_type="Uterus, leiomyosarcoma"
     /clone_lib="NIH_MGC_71"
     /lab_host="DH10B"
     /note="Vector: PCMV-SPORT6"
26..493
     /codon_start=1
     /product="Unknown (protein for MGC:8764)"
     /protein_id="AAH15013.1"
     /db_xref="GI:15929110"
     /translation="MEIRSGVSGQAVARRMDGSDRGGGKDKATGSDYENLPTSA
VLTGVNVMIMAGGAPAHAMFYACENMKRTLNDVFHHQGNHSLANGILKAFYWS"
BASE COUNT      215 a   220 c   231 g   192 t
ORIGIN

Query Match      29.6%; Score 172; DB 9; Length 858;
Best Local Similarity 69.1%; Pred. No. 2e-18;
Matches 235; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 183 ggactacaggcgctgcggctggagcaactgtcaccagcacatgggtggcaggcgct 242
Db 127 GGACTACGAGAACCGACCTGAGCGCTTCCTGCTGCCACCCACATGACAGCAGGACGAT 186

QY 243 ggcaggatctctggagcactgcgtgatgtaccccatcgactcgctcaagaccggatgca 302
Db 187 GCCCGGATCTCGGAGCACTCGTATGTACCCGGTGACTCGGTCAAGACACGAATGCA 246

QY 303 gagtctacagctgaccagctgcgcgcgtatcgcaatgtgttggaggccctctggagat 362
Db 247 GAGTTTGTAGTCCAGATCCCAAGGCCAGTACACAAAGTATCTACGGAGCCCTCAAGAAAT 306

QY 363 tataaagcaggagcctatgagcccatgaggggctgaacgtcacagcaacaggcgc 422
Db 307 CATGGCGACCGAAGCGCTTCTGGAGGCCCTTGCGAGCGGCTCAACGTCATGATCGGTGC 366

QY 423 agggcctgcgcacgcctcttttgcctgtcacgaaagttaaaagacattgattga 482
Db 367 AGGGCGGCCCATGCCTGCTATTTTGCCTGCTATGAAACATGAAAGACGCTTTAAATGA 426

QY 483 tgtaatccacctggggcaatagccatttgcctgaatggt 522
Db 427 CGTTTTCCACCACCAAGAAACAGCACCTAGCCCAACGCT 466

RESULT 15
AC099368
LOCUS
DEFINITION      Rattus norvegicus clone CH230-84G11, *** SEQUENCING IN PROGRESS
ACCESSION      AC099368
VERSION        AC099368.2  Gi:17973175
KEYWORDS       HTG; HTGS_PHASE1.

SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 163205)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brileva,M., Brown,E., Brown,N.P., Bryant,C., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenko,S.,
Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojebokan,I., Rolfe,W.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,K., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 163205)
Worley,K.C.
Direct Submission
Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16901901.
-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Center project name: GHJ
Center clone name: CH230-84G11
-----
Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrap!
Consensus quality: 127058 bases at least Q40
Consensus quality: 138264 bases at least Q30
Consensus quality: 146548 bases at least Q20
Estimated insert size: 120308; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length

```

*(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a "working draft" sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 8509: contig of 8509 bp in length
* 8510 8609: gap of unknown length
* 8610 13540: contig of 4931 bp in length
* 13541 13640: gap of unknown length
* 13641 19516: contig of 5876 bp in length
* 19517 19616: gap of unknown length
* 19617 24990: contig of 5374 bp in length
* 24991 25090: gap of unknown length
* 25091 29401: contig of 4311 bp in length
* 29402 29501: gap of unknown length
* 29502 33053: contig of 3552 bp in length
* 33054 33153: gap of unknown length
* 33154 38439: contig of 5286 bp in length
* 38440 38539: gap of unknown length
* 38540 41183: contig of 2644 bp in length
* 41184 41283: gap of unknown length
* 41284 44936: contig of 3653 bp in length
* 44937 45036: gap of unknown length
* 45037 47637: contig of 2601 bp in length
* 47638 47737: gap of unknown length
* 47739 50166: contig of 2429 bp in length
* 50167 50266: gap of unknown length
* 50267 54170: contig of 3904 bp in length
* 54171 54270: gap of unknown length
* 54271 57672: contig of 3402 bp in length
* 57673 57772: gap of unknown length
* 57773 61162: contig of 3390 bp in length
* 61163 61262: gap of unknown length
* 61263 63890: contig of 2628 bp in length
* 63891 63990: gap of unknown length
* 63991 66900: contig of 2910 bp in length
* 66901 67000: gap of unknown length
* 67001 69817: contig of 2717 bp in length
* 69818 69817: gap of unknown length
* 69818 72921: contig of 3104 bp in length
* 72922 73021: gap of unknown length
* 73022 75146: contig of 2125 bp in length
* 75147 75246: gap of unknown length
* 75247 79038: contig of 3792 bp in length
* 79039 79138: gap of unknown length
* 79139 80340: contig of 1202 bp in length
* 80341 80440: gap of unknown length
* 80441 82366: contig of 1926 bp in length
* 82367 82466: gap of unknown length
* 82467 84184: contig of 1718 bp in length
* 84185 84284: gap of unknown length
* 84285 86001: contig of 1717 bp in length
* 86002 86101: gap of unknown length
* 86102 87695: contig of 1594 bp in length
* 87696 87795: gap of unknown length
* 87796 89287: contig of 1492 bp in length
* 89288 89387: gap of unknown length
* 89388 91950: contig of 2563 bp in length
* 91951 92050: gap of unknown length
* 92051 94575: contig of 2525 bp in length
* 94576 94675: gap of unknown length
* 94676 96581: contig of 1906 bp in length
* 96582 96681: gap of unknown length
* 96682 98235: contig of 1554 bp in length
* 98236 98335: gap of unknown length
* 98336 99814: contig of 1479 bp in length
* 99815 99914: gap of unknown length
* 99915 100964: contig of 1050 bp in length
* 101064: gap of unknown length

* 101065 103074: contig of 2010 bp in length
* 103075 103174: gap of unknown length
* 103175 104311: contig of 1137 bp in length
* 104312 104411: gap of unknown length
* 104412 106834: contig of 2423 bp in length
* 106835 108934: gap of unknown length
* 108935 108922: contig of 1988 bp in length
* 108923 108922: gap of unknown length
* 109023 110868: contig of 1846 bp in length
* 110869 110968: gap of unknown length
* 110969 113056: contig of 2088 bp in length
* 113057 113156: gap of unknown length
* 113157 114296: contig of 1140 bp in length
* 114297 114396: gap of unknown length
* 114397 116312: contig of 1916 bp in length
* 116313 116412: gap of unknown length
* 116413 117492: contig of 1080 bp in length
* 117493 117592: gap of unknown length
* 117593 119552: contig of 1960 bp in length
* 119553 119652: gap of unknown length
* 119653 120658: contig of 1006 bp in length
* 120659 120758: gap of unknown length
* 120759 121895: contig of 1137 bp in length
* 121896 121995: gap of unknown length
* 121996 123272: contig of 1277 bp in length
* 123273 123372: gap of unknown length
* 123373 125310: contig of 1938 bp in length
* 125311 125410: gap of unknown length
* 125411 126891: contig of 1481 bp in length
* 126892 126991: gap of unknown length
* 126992 128690: contig of 1699 bp in length
* 128691 128790: gap of unknown length
* 128791 130827: contig of 2037 bp in length
* 130828 130927: gap of unknown length
* 130928 132501: contig of 1574 bp in length
* 132502 132601: gap of unknown length
* 132602 134071: contig of 1470 bp in length
* 134072 134171: gap of unknown length
* 134172 135838: contig of 1667 bp in length
* 135839 135938: gap of unknown length
* 135939 137589: contig of 1651 bp in length
* 137590 137689: gap of unknown length
* 137690 139043: contig of 1354 bp in length
* 139044 139143: gap of unknown length
* 139144 140777: contig of 1634 bp in length
* 140778 140877: gap of unknown length
* 140878 142526: contig of 1649 bp in length
* 142527 142626: gap of unknown length

Query Match 21.5% Score 125.2; DB 2; Length 163205;

Best Local Similarity 87.4%; Pred. No. 2.6e-11;

Matches 194; Conservative 0; Mismatches 23; Indels 5; Gaps 5;

Qy 304 agtctacagcctgaccag-cgtgccgcctatcgcaatgtgttga-ggccccttgagga 361
Db 13720 ATTTTACAGCCTGATCCAGCGCCCGTTATCGAAAAAGTTGTGAGGCTTTTGGAGAA 13779
Qy 362 ttataagaacggaggcctatggagggccatga-gggggctgaacgtcacagcaagc 420
Db 13780 TTATAAGAACAAGAGCGCCCTGTGGAGGCCCATGAGGGGGTTGAACCTCAGTCCACTGGC 13839
Qy 421 gcaggcctgccacgccttatttgcctgctacgaaagttaaaaaagacattga 480
Db 13840 GCGGGGCTGCCACAGCCCTCTATTG-CYGCCTACGAAAAG-TAAAAAGACATTGAGT 13897
Qy 481 gatgtaaccaccctgggggcaatagccattatgccaatggt 522
Db 13998 GATCTAATCCACCCGCGGGGCAATAGCCATATTGCCAATGCT 13939

Search completed: September 28, 2002, 07:53:07
Job time: 14508 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: September 28, 2002, 04:53:25 ; Search time 664.31 Seconds
(without alignments)
1504.184 Million cell updates/sec

Title: US-09-870-113-3

Perfect score: 582

Sequence: 1 atgaggttgaggggcgggg.....tgaacctgcggaaggctga 582

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

```

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	469.4	80.7	1336	21 AAC90457	Human uncoupling p
2	462	79.4	1418	21 AAC74843	Human ORFX ORF398
3	454.4	78.1	1322	22 AAF27733	Human transport pr
4	440	75.6	1294	22 AAI60661	Human polynucleoti
5	401	68.9	2502	22 AAS03906	Human secreted pro
6	398.6	68.5	1316	22 AAI58875	Human polynucleoti
7	273	45.9	401	22 AAF66156	Novel human polynu
8	233	40.0	27960	22 AAK69779	Human immune/haema
9	233	40.0	27960	22 AAK73320	Human immune/haema

10	200.4	34.4	1902	22	AAK52172	Human polynucleoti
11	196.6	33.8	1305	22	AAF59920	Human mitochondria
12	187.4	32.2	2562	22	AAH90057	Human bone marrow
13	181	31.1	339	22	AAK58672	Human immune/haema
14	174.2	29.9	1890	22	AAK94897	Human full-length
15	173.6	29.8	842	22	AAK92357	Human CDNA 5'-end
16	173.6	29.8	842	22	AAK93895	Human CDNA clone r
17	173.6	29.8	1814	22	AAK94871	Human full-length
18	172.6	29.7	1468	21	AAK77175	Human ORFX ORF2730
19	172.6	29.6	3028	22	AAK91320	Human polynucleoti
20	170.4	29.3	622	22	AAK91764	Human CDNA 5'-end
21	170.4	29.3	622	22	AAK93200	Human CDNA clone r
22	168.6	29.0	1913	22	AAH89944	Human bone marrow
23	160	27.5	1758	20	AAV84573	Human secreted pro
24	160	27.5	1758	22	ABA83356	Human secreted pro
25	132.8	22.8	1716	21	AAK77189	Human ORFX ORF2744
26	121	20.8	300	20	AAZ13026	Human gene express
27	121	20.8	710	20	AAZ15876	Human gene express
28	109.8	18.9	2837	22	AAK80493	Human immune/haema
29	105.2	18.1	483	24	AB199871	Mouse ischaemic co
30	79.8	13.7	673	21	AAF12614	Aspergillus oryzae
31	72.4	12.4	2137	23	ABL08627	Drosophila melanog
32	67.8	11.6	3058	22	AAK83876	Human immune/haema
33	65.8	11.3	448	22	AAK77174	Human ORFX ORF2729
34	63.2	10.9	1000	21	AAA02484	Human colon cancer
35	59.2	10.2	349	22	AAI84504	Human polynucleoti
36	58.2	10.0	318	21	AAA38184	Primer used in the
37	58.2	10.0	320	21	AAA38186	Primer used in the
38	58.2	10.0	2188	20	AAZ77506	Human ovarian tumo
39	58.2	10.0	3198	20	AAK02974	Human IL-1ra BAC c
40	58	10.0	2037	24	AAK562787	CDNA sequence #574
41	57.6	9.9	320	21	AAA38185	Primer used in the
42	57.6	9.9	1698	23	AAK84408	DNA encoding novel
43	57.6	9.9	2049	23	AAK76210	DNA encoding novel
44	57.4	9.9	320	21	AAA38183	Primer used in the
45	57.2	9.8	1286	24	AB199656	Mouse ischaemic co

ALIGNMENTS

RESULT 1
AAC90457
ID AAC90457 standard; CDNA; 1336 BP.

XX AAC90457;

XX 12-MAR-2001 (first entry)

DE Human uncoupling protein cDNA #6.

XX Human; uncoupling protein; immunosuppressive; antiarthritic;

KW antirheumatic; antiproliferative; cardiant; vasotropic;

KW cerebroprotective; neuroprotective; antibacterial; ophthalmological;

KW gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;

KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;

KW infertility; ss.

XX Homo sapiens.

XX WO200061614-A2.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US09534.

XX 09-APR-1999; 99US-0128701.

XX 08-JUL-1999; 99US-0142821.

XX 18-AUG-1999; 99US-0149448.

XX 12-NOV-1999; 99US-0164751.

XX (HUMA-) HUMAN GENOME SCI INC.

XX

Db 61 cgggactacgagcgtgcggctgagccactgtcaccacgcacatggtggcagcgcc 120
 QY 241 gtggcaggatccttgagcactgcgtgatgtaccacatcgactgcgtcaagaccggtg 300
 Db 121 gtggcaggatccttgagcactgcgtgatgtaccacatcgactgcgtcaagaccggtg 180
 QY 301 cagagtctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctggagg 360
 Db 181 cagagtctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctggagg 240
 QY 361 attataaagcaggagggcctatggagcccatgaggggctgaaagctcacagcaacaggc 420
 Db 241 attataaagcaggagggcctatggagcccatgaggggctgaaagctcacagcaacaggc 300
 QY 421 gaggggcctgcccagcgcctttatttgcctgctacgaaaaagttaaaaaagacattgagt 480
 Db 301 gaggggcctgcccagcgcctttatttgcctgctacgaaaaagttaaaaaagacattgagt 360
 QY 481 gatgtaataccaccctgggggcaatagccatattgccaatggtgcggcgggtgtgtggca 540
 Db 361 gatgtaataccaccctgggggcaatagccatattgccaatggtgcggcgggtgtgtggca 420
 QY 541 acattacttcattgatgcagccatgaaccctgcgggaagcctga 582
 Db 421 acattacttcattgatgcagccatgaaccctgcgggaagcctga 462

RESULT 3

AAAF27733 standard; cDNA: 1322 BP.

XX AC AAF27733;

DT 28-MAR-2001 (first entry)

DE Human transport protein TPPT-33 coding sequence.

KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
 KW neurological disorder; cardiovascular disorder; reproductive disorder;
 KW immune disorder; cancer; ss.

OS Homo sapiens.

XX PN WO200078953-A2.

XX PD 28-DEC-2000.

XX PF 16-JUN-2000; 2000WO-US16668.

XX PR 17-JUN-1999; 99US-0139923.

XX PR 10-AUG-1999; 99US-0148177.

XX PR 18-AUG-1999; 99US-0149357.

XX PR 28-OCT-1999; 99US-0162287.

XX FA (INCY-) INCYTE GENOMICS INC.

PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
 PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;

XX DR WPI; 2001-041424/05.

XX DR P-PSDB; AAB60113.

XX Isolated polypeptide with a human transport protein sequence is useful
 XX for the diagnosis, prevention and treatment of disorders associated
 XX with the immune, reproductive and cardiovascular systems -

PS Claim 5; Page 160; 165pp; English.

XX The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPPTs). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative

CC disorders such as cancer.

XX SQ Sequence 1322 BP; 328 A; 376 C; 351 G; 267 T; 0 other;
 Matches 458; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Query Match 78.1%; Score 454.4; DB 22; Length 1322;
 Best Local Similarity 98.7%; Pred. NO. 2e-82;
 Matches 458; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 114 cggggggccggcggcgggagggccggtgagggccggtacgacaaagatcccgga 173
 Db 2 cggggggccggcggcgggagggccggtgagggccggtacgacaaagatcccgga 61
 QY 174 ctccggcccgactacgagggcgtccggtgagccactgtcacccagcacatggtggc 233
 Db 62 ctccggcccgactacgagggcgtccggtgagccactgtcacccagcacatggtggc 121
 QY 234 aggcgcctgagggatctctggagcactggtgtatcccccactgcactgcgtcgaagac 293
 Db 122 aggcgcctgagggatctctggagcactggtgtatcccccactgcactgcgtcgaagac 181
 QY 294 cgggatgcagagctcacagcctgaccagctgacccagctgcccgcgtatcgcaatggtggagccct 353
 Db 182 cgggatgcagagctcacagcctgaccagctgacccagctgcccgcgtatcgcaatggtggagccct 241
 QY 354 ctggaggattataaagacggaggcctatgagggccctgagggcctgaggggctgaacctcacagc 413
 Db 242 ctggaggattataaagacggaggcctatgagggcctgagggcctgaggggctgaacctcacagc 301
 QY 414 aacaggcggcggcgtcccgccctttatttgcctgctacgaaaaagttaaaaaagac 473
 Db 302 aacaggcggcggcgtcccgccctttatttgcctgctacgaaaaagttaaaaaagac 361
 QY 474 attgagtgtatgaatccaccctgggggcaatagccatattgccaatggtgcggcgggtg 533
 Db 362 attgagtgtatgaatccaccctgggggcaatagccatattgccaatggtgcggcgggtg 421
 QY 534 tgtggcaacattacttcattgatgcagccatgaacctgcggaag 577
 Db 422 tgtggcaacattacttcattgatgcagccatgaacctgcggaag 465

RESULT 4

AAI60661/c standard; cDNA: 1294 BP.

XX AC AAI60661;

DT 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 4650.

XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

[illegible]

|||||
Db 66 ccggactcggccggactacgagcgctgccggtggagccactgtcaccacgcacatg 125
Qy 229 gfggcaggcgcgtgagcggatccctgagcactgcgtgatgtacccatcgaactgcgtc 288
Db 126 gfggcaggcgcgtgagcggatccctgagcactgcgtgatgtacccatcgaactgcgtc 185
Qy 289 aagaccggatgcagagtcacagctgaccagctgcccgctatccgaatgtgtggag 348
Db 186 aagaccggatgcagagtcacagctgaccagctgcccgctatccgaatgtgtggag 245
Qy 349 gccctctgagatataagaacgagggcctctatggagcccatgaggggctgaacgtc 408
Db 246 gccctctgagatataagaacgagggcctctatggagcccatgaggggctgaacgtc 305
Qy 409 acagcaacagcgcgagggcctccacgccttttttttttttttttttttttttttttt 468
Db 306 acagcaacagcgcgagggcctccacgccttttttttttttttttttttttttttttt 365
Qy 469 aagacattgagtgatgaatccaccctgggggcaatagccatattgccaatggt 522
Db 366 aagacattgagtgatgaatccaccctgggggcaatagccatattgccaatggt 419
RESULT 6
AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX
AC AAI58875;
XX
XX
Df 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1078.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0486725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39719.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 1078; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

the encoded polypeptides (AAM38642-AAM42213) with nootropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening,
assays for receptor activity, arthritis and inflammation, leukaemias and
C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;
Query Match 68.5%; Score 398.6; DB 22; Length 1316;
Best Local Similarity 95.5%; Pred. No. 3.1e-71;
Matches 421; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
Qy 138 cggggcctgagggcccggtacgacaaagtcggactccgcccggactacgagcgct 197
Db 15 cgaggcctgagggcccggtacgacaaagtcggactccgcccggactacgagcgct 74
Qy 198 gccggctgagggcactgtcaccacgcacatg-gtggcaggcgccgtggcaggatcctgg 256
Db 75 gtcggctgagggcactgtcaccacgcacatg-gtggcaggcgccgtggatgctg 134
Qy 257 agcactgcgtgatgtaccccatcgaactcgtcaagaccgcgcatgcagagtctacagcctg 316
Db 135 agcactgcgtgatgtaccccatcgaactcgtcgtgatgcccgcatgcatgtctacagctg 194
Qy 317 acccagctgccgctatcgcaatgtgttggaggccctctggagattataagaacggagg 376
Db 195 acccagctgccgctatcgcaatgtgttggaggctcctctggagattataagaacggagg 254
Qy 377 gcctatgagggccatgaggggctgaactcagacagacagggcgagggcctgccacg 436
Db 255 gcctatgagggccatgaggggctgaactcagacagggcgagggcctgccacg 314
Qy 437 ccccttatttgcctgtacgaaaagttaaaaagacacattgagtgatgaatccaccctg 496
Db 315 ccccttatttgcctgtacgaaaagttaaaaagacacattgagtgatgaatccaccctg 374
Qy 497 ggggcaatagccatattgccaatggtgcggcggggtgtgtggcaacattacttcattg 556
Db 375 ggggcaatagccatattgccaatggtgcggcggggtgtgtggcaacattacttcattg 434
Qy 557 cagccatgaacccctgcggaag 577
Db 435 cagccatgaacccctgcggaag 455
RESULT 7
AAF66156
ID AAF66156 standard; cDNA; 401 BP.
XX
AC AAF66156;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1912.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX

PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
XX
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Iamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
XX WPI; 2001-091805/10.
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
XX Claim 9; Page 818; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match 46.9%; Score 273; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 4.3e-46;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 gtctacagcctgacccagctgcccgtatcgcaatgtgttgaggccctctgaggatta 364
Db 9 gtctacagcctgacccagctgcccgtatcgcaatgtgttgaggccctctgaggatta 68

Qy 365 taagaacgaggggcctatgagggccatgagggtgaaacgtcacagcaacagggcgag 424
Db 69 taagaacgaggggcctatgagggccatgagggtgaaacgtcacagcaacagggcgag 128

Qy 425 ggcctgccacgccccttatttgcctgctacgaaagttaaaagacattgagtgatg 484
Db 129 ggcctgccacgccccttatttgcctgctacgaaagttaaaagacattgagtgatg 188

Qy 485 taatccacctgggggcaatagccatattccaatgtgcgcgggtgtgtgccaacat 544
Db 189 taatccacctgggggcaatagccatattccaatgtgcgcgggtgtgtgccaacat 248

Qy 545 tacttcgatgagccatgaacctgcggaag 577
Db 249 tacttcgatgagccatgaacctgcggaag 281

RESULT 8
AAK69779
ID AAK69779 standard; DNA; 27960 BP.
XX
AC AAK69779;
XX

DT 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24591.
DE
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
OS
OS Homo sapiens.
PN
PN WO200157182-A2.
XX
XX 09-AUG-2001.
PD
PD
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR
PR 04-FEB-2000; 2000US-0180628.
PR
PR 24-FEB-2000; 2000US-0184664.
PR
PR 02-MAR-2000; 2000US-0186350.
PR
PR 16-MAR-2000; 2000US-0189874.
PR
PR 17-MAR-2000; 2000US-0190076.
PR
PR 18-APR-2000; 2000US-0198123.
PR
PR 19-MAY-2000; 2000US-0205515.
PR
PR 07-JUN-2000; 2000US-0209467.
PR
PR 28-JUN-2000; 2000US-0214886.
PR
PR 30-JUN-2000; 2000US-0215135.
PR
PR 07-JUL-2000; 2000US-0216647.
PR
PR 11-JUL-2000; 2000US-0216880.
PR
PR 11-JUL-2000; 2000US-0217487.
PR
PR 11-JUL-2000; 2000US-0217496.
PR
PR 14-JUL-2000; 2000US-0218290.
PR
PR 26-JUL-2000; 2000US-0220963.
PR
PR 26-JUL-2000; 2000US-0220964.
PR
PR 14-AUG-2000; 2000US-0224518.
PR
PR 14-AUG-2000; 2000US-0224519.
PR
PR 14-AUG-2000; 2000US-0225213.
PR
PR 14-AUG-2000; 2000US-0225214.
PR
PR 14-AUG-2000; 2000US-0225266.
PR
PR 14-AUG-2000; 2000US-0225267.
PR
PR 14-AUG-2000; 2000US-0225268.
PR
PR 14-AUG-2000; 2000US-0225270.
PR
PR 14-AUG-2000; 2000US-0225447.
PR
PR 14-AUG-2000; 2000US-0225757.
PR
PR 14-AUG-2000; 2000US-0225758.
PR
PR 14-AUG-2000; 2000US-0225759.
PR
PR 18-AUG-2000; 2000US-0226279.
PR
PR 22-AUG-2000; 2000US-0226681.
PR
PR 22-AUG-2000; 2000US-0227182.
PR
PR 23-AUG-2000; 2000US-0227009.
PR
PR 30-AUG-2000; 2000US-0228924.
PR
PR 01-SEP-2000; 2000US-0229287.
PR
PR 01-SEP-2000; 2000US-0229343.
PR
PR 01-SEP-2000; 2000US-0229344.
PR
PR 01-SEP-2000; 2000US-0229345.
PR
PR 05-SEP-2000; 2000US-0229509.
PR
PR 05-SEP-2000; 2000US-0229513.
PR
PR 06-SEP-2000; 2000US-0230437.
PR
PR 06-SEP-2000; 2000US-0230438.
PR
PR 08-SEP-2000; 2000US-0231242.
PR
PR 08-SEP-2000; 2000US-0231243.
PR
PR 08-SEP-2000; 2000US-0231244.
PR
PR 08-SEP-2000; 2000US-0231413.
PR
PR 08-SEP-2000; 2000US-0231414.
PR
PR 08-SEP-2000; 2000US-0232080.
PR
PR 08-SEP-2000; 2000US-0232081.
PR
PR 12-SEP-2000; 2000US-0231968.
PR
PR 14-SEP-2000; 2000US-0232397.
PR
PR 14-SEP-2000; 2000US-0232398.
PR
PR 14-SEP-2000; 2000US-0232399.
PR
PR 14-SEP-2000; 2000US-0232400.
PR
PR 14-SEP-2000; 2000US-0232401.
PR
PR 14-SEP-2000; 2000US-0233063.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX WO200157182-A2.
PN
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.


```

Db 1280 gcagagtttgatccagatcccaagccaggtacacaagtgtctacggagccctcaagaa 1339
Qy 360 gattataaagagcggcctatgagcccatgagggcctgagggcctgacagtcacagcaacagg 419
Db 1340 aatcatcg9gacccgaagcctctctg9gagcccttgc9aggcctcaacgtcatgcatggg 1399
Qy 420 cgcaggcctgcccacgcccctttatttgcctgctacgaaagttaaaaaagacattggag 479
Db 1400 tgcaggccagcccatgccatgtatttgcctgctatgaaacatgaaaggactttaaa 1459
Qy 480 tgaatgaatcaacccctggggccaatagccatattgccaatggtgcggcgggtgtgtggc 539
Db 1460 tgacgttttcaccacccaagaaacagaccactagccaacggatagctgggagtatggc 1519
Qy 540 aacattactcatgatgcagccatgacgaacctgcggaag 577
Db 1520 caccctgtccacgatgcggttaatgaatccagcagaag 1557

RESULT 11
AAF59920
ID AAF59920 standard; cdna; 1305 BP.
XX
AC AAF59920;
XX
DT 04-MAY-2001 (first entry)
XX
DE Human mitochondrial solute carrier protein hMSC-o cDNA.
XX
DE Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
KW preparation; detection; ss.
KW
XX Homo sapiens.
XX
XX CN1269409-A.
XX
PD 11-OCT-2000.
XX
PF 17-MAR-2000; 2000CN-0114958.
XX
PR 17-MAR-2000; 2000CN-0114958.
XX
PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
XX
PI Zhang X, Gao X, Xiao H;
XX
DR WPI; 2001-050544/07.
DR P-PSDB; AAB60658.
XX
New human mitochondrion solute carrier protein and its nucleic acid -
PS Claim 1; Page 19-20; 21pp; Chinese.
XX
CC The invention relates to a novel human mitochondrial solute carrier
CC protein, hMSC-o (AAB60658), and cDNA encoding it (AAF59920). hMSC-o is
CC expressed in normal human hypothalamus tissue. The invention also relates
CC to the preparation of hMSC-o proteins and nucleic acids, and the
CC detection of hMSC-o proteins and nucleic acids in a sample. The present
CC sequence represents cDNA encoding hMSC-o.
XX
SQ Sequence 1305 BP; 292 A; 388 C; 363 G; 262 T; 0 other;

Query Match 33.88; Score 196.6; DB 22; Length 1305;
Best Local Similarity 68.68; Pred. No. 9.5e-31;
Matches 271; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 183 ggactacagagcgtgcgcgctgagccactgtcaccacgacacatggtggcagcgccgt 242
Db 57 ggactacagaaactgcgactagcgtccctcgttccaccacatgacagcagcgat 116
Qy 243 ggcaggatctctggagcactgcgtgatgtacctccatgcgtcgtcagaccggatgca 302

```

```

Db 117 ggcgggatacctcgagcactcggtcattcaccggtgactcggtagaacacacgaatgca 176
Qy 303 gagtctacagcctgaccagcgtgcccgctatcgcgaatgtgttgaggccctctggaggat 362
Db 177 gagtttgagtcacgatacccaagccaggtacacaagtatctacggagccctcaagaaat 236
Qy 363 tataagaacggaggcctatgagggcccatgagggcctgagggcgtcacagcaacaggcgc 422
Db 237 catcggaaccgaagccttctgaggcccttgcgaggcgtcaacgtcatgatcatgggtgc 296
Qy 423 aggcctgcccacgcccctttatttgcctgctacgaaagttaaaaaagacattgagtga 482
Db 297 aggcgcggcccatgccatgtatttgcctgctatgaaacatgaaaggactttaaaatga 356
Qy 483 tgtaatccacccctggggccaatagccatattgccaatggtgcggcgggtgtgtggcaac 542
Db 357 cgtttccaccaccaagaaacagccacctagccaacgggagatagctgggagtatggccac 416
Qy 543 attactcatgatgcagccatgaaacctgcggaag 577
Db 417 cctgctccacgatgcggttaatgaatccagcagaag 451

RESULT 12
AAH90057
ID AAH90057 standard; cdna; 2562 BP.
XX
AC AAH90057;
XX
DT 01-OCT-2001 (first entry)
XX
DE Human bone marrow cDNA, SEQ ID NO: 301.
XX
DE
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
XX Homo sapiens.
XX
XX WO200153453-A2.
XX
PD 26-JUL-2001.
XX
PF 23-DEC-2000; 2000WO-US34960.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Drmanac RT;
XX
XX WPI; 2001-488707/53.
XX P-PSDB; AAM00938.
XX
Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1; Page 412-413; 648pp; English.
XX
XX The present sequence is one of 251 novel human polynucleotides
CC expressed in the bone marrow. The polynucleotide and the

```


PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-483426/52.
DR
P-PSDB; AAM85891.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis.
XX
XX Claim 1; SEQ ID NO 3732; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 339 BP; 86 A; 84 C; 101 G; 64 T; 4 other;

Query Match 31.1%; Score 181; DB 22; Length 339;
Best Local Similarity 95.1%; Pred. No. 1.2e-27;
Matches 195; Conservative 1; Mismatches 8; Indels 1; Gaps 1;
Qy 291 gaccggagtcagagctctacagcctgaccagctgccgctatcgcaatgtgttgaggc 350
|||||
Db 136 gaccggagtcagagctctacagcctgaccagctgccgctatcgcaatgtgttgaggc 195
Qy 351 cctctggaggattataagaacggaggcgctatggaggccctatggaggctgaacgtcac 410
|||||
Db 196 cctctggaggattataagaacggaggcgctatggagg-ccatgaggggctgaacgtcac 254
Qy 411 agcaacaggcgaggcgctgccacgcctttattttgcctgctacgaaagttaaaaaa 470
|||||
Db 255 agcaacaggcgaggcgctnccacgcctttattttgcctgctacgaaagttaaaaaa 314
Qy 471 gacattgagtgatgaatccaccct 495
Db 315 ggcattgagtgatgaatccaccct 339

RESULT 14
AAK94897
ID AAK94897 standard; cDNA; 1890 BP.
XX
XX AAK94897;
AC
XX
XX 06-NOV-2001 (first entry)
DT Human full-length cDNA, SEQ ID NO: 4109.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
KW Homo sapiens.
XX
XX EP1130094-A2.
PN
XX
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-0114089.
PF
XX
XX 08-JUL-1999; 99JP-0194486.
PR
XX
XX 11-JAN-2000; 2000JP-0118774.
PR
XX
XX 02-MAY-2000; 2000JP-0183765.
PR
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 04:48:35 ; Search time 182.06 Seconds
(without alignments)
785,228 Million cell updates/sec

Title: US-09-870-113-3
Perfect score: 582
Sequence: 1 atgagttgagggcgggg.....tgaacctcggaaggtga 582

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62.4	10.7	319	4 US-09-165-264-8	Sequence 8, Appli
2	58.2	10.0	318	4 US-09-165-264-12	Sequence 12, Appl
3	58.2	10.0	320	4 US-09-165-264-14	Sequence 14, Appl
c 4	58.2	10.0	152331	3 US-09-128-155-16	Sequence 16, Appl
5	57.8	9.9	320	4 US-09-165-264-7	Sequence 7, Appli
6	57.6	9.9	320	4 US-09-165-264-13	Sequence 13, Appl
7	57.4	9.9	320	4 US-09-165-264-11	Sequence 11, Appl
c 8	50.6	8.7	833	2 US-08-403-852D-3	Sequence 3, Appli
c 9	50.6	8.7	833	2 US-08-510-646B-3	Sequence 3, Appli
c 10	50.6	8.7	833	4 US-09-231-818-3	Sequence 3, Appli
c 11	50.6	8.7	5392	2 US-08-403-852D-1	Sequence 1, Appli
c 12	50.6	8.7	5392	3 US-08-510-646B-1	Sequence 1, Appli
c 13	50.6	8.7	5392	4 US-09-231-818-1	Sequence 1, Appli
c 14	49.4	8.5	43280	2 US-08-804-277C-1	Sequence 1, Appli
15	47	8.1	1227	2 US-08-903-800A-1	Sequence 1, Appli
c 16	46.8	8.0	1931	2 US-09-130-114-2	Sequence 2, Appli
c 17	46.8	8.0	4257	2 US-08-690-473-1	Sequence 1, Appli
c 18	46.8	8.0	4257	4 US-09-259-821A-1	Sequence 1, Appli
c 19	46.8	8.0	4257	4 US-08-843-659-1	Sequence 1, Appli
c 20	45.8	8.0	12001	1 US-08-458-568A-11	Sequence 11, Appl
21	45.6	8.0	4403765	4 US-09-103-840A-2	Sequence 2, Appli
22	46.6	8.0	4411529	4 US-09-103-840A-1	Sequence 1, Appli
c 23	46.2	7.9	2823	2 US-08-398-008A-1	Sequence 1, Appli
c 24	46.2	7.9	2823	2 US-08-893-333-1	Sequence 1, Appli
25	45.2	7.9	3415	1 US-08-054-077C-1	Sequence 1, Appli
26	45.8	7.9	460	2 US-08-903-800A-4	Sequence 4, Appli
c 27	45.6	7.8	1300	4 US-08-483-533-39	Sequence 39, Appli

c 28	45.6	7.8	1300	4 US-09-283-471A-39	Sequence 39, Appl
c 29	45.6	7.8	53526	3 US-08-658-136-2	Sequence 2, Appli
c 30	45.6	7.8	53577	3 US-08-658-136-1	Sequence 1, Appli
c 31	45.4	7.8	1327	4 US-08-483-533-36	Sequence 36, Appl
c 32	45.4	7.8	1327	4 US-09-283-471A-36	Sequence 36, Appl
c 33	45.4	7.8	1335	5 PCT-US91-06532-1	Sequence 1, Appli
34	45.2	7.8	1028	4 US-08-118-200-1	Sequence 1, Appli
35	45.2	7.8	1028	4 US-08-458-745-1	Sequence 1, Appli
c 36	45.2	7.8	12001	1 US-08-458-568A-11	Sequence 11, Appl
37	45	7.7	8438	1 US-07-945-283-1	Sequence 1, Appli
c 38	44.6	7.7	5228	4 US-09-428-711A-15	Sequence 15, Appl
39	44.6	7.7	23673	4 US-09-773-816-1	Sequence 1, Appli
40	44.4	7.6	405	2 US-08-903-800A-3	Sequence 3, Appli
41	44.2	7.6	700	4 US-09-236-097-7	Sequence 7, Appli
42	44.2	7.6	1026	4 US-07-751-891B-24	Sequence 24, Appl
c 43	44.2	7.6	1868	1 US-08-658-883B-1	Sequence 1, Appli
44	44.2	7.6	4488	4 US-08-406-030A-3	Sequence 3, Appli
45	44.2	7.6	5222	4 US-07-751-891B-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-165-264-8
; Sequence 8, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamorthy, Thuraiyayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-8

Query Match	10.7%;	Score 62.4;	DB 4;	Length 319;
Best Local Similarity	63.2%;	Pred. No. 8.5e-06;		
Matches	96;	Conservative	0;	Mismatches 56; Indels 0; Gaps 0;
QY	3	gagtttgagggcggggtgctgctgagcggtgtgagggggcgagggcgagggcgagggcg 62		
Db	161	gg 220		
QY	63	gagccccgggagtcgagcgctgagcggtgagcgagcgagggcgagggcgagggcg 122		
Db	221	gg 280		
QY	123	cggcgcgagggcgagggcgagggcgagggcgagggcgagggcgagggcgagggcg 154		
Db	281	gg 312		

RESULT 2
US-09-165-264-12
; Sequence 12, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamorthy, Thuraiyayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1


```
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/231.818
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403.852
; FILING DATE: 10-MAY-1995
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: S.pristinaespiralis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..833
; US-09-231-818-3

Query Match      8.7%; Score 50.6; DB 4; Length 833;
Best Local Similarity 53.0%; Pred. No. 0.0043;
Matches 133; Conservative 0; Mismatches 114; Indels 4; Gaps 1;

Qy 3 ggaattgagggcggggtgctgcggtgtgctggggggcgccgagggcccgccggcg 62
Db 787 GCGGTTCGCGCAGGGTGTGTGCGCGCGGTGCGGTCCGGCGCGCGCGCGCGCGCG 728

Qy 63 gagcc-----ccggggagtcggcgctgctggagcgggtggtgctgagcgggcggtg 118
Db 727 CGGCCAGGACCGGTCTGTCGGAGCGCGGTGAGCGCCACCGTGAAGCCGTCGCGGGCTGT 668

Qy 119 gggccggcgggcgggagggcgccggcgccgctgagcccccgggtacgaagatccggactccg 178
Db 667 CGCGCGGCGAGTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 608

Qy 179 gcccgggactagagggcgctgcggcgctgagccactgtccaccacacacatgtgcagggc 238
Db 607 GCTGCTTACCGAGGCGCGCTGCGCGCGCTGCGCGCGCTGCTGAGCAGGACGAGGTTCGCGCGCG 548

Qy 239 ccgtggcaggg 249
Db 547 CGAGGCGCGGG 537

RESULT 11
us-08-403-852D-1/c
; Sequence 1, Application US/08403852D
; Patent No. 5891695
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
```

```
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; Biosynthesis Of Streptogramins, Nucleotide Sequences
; Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/403.852D
; FILING DATE: 10-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: S.pristinaespiralis
; US-08-403-852D-1

Query Match      8.7%; Score 50.6; DB 2; Length 5392;
Best Local Similarity 53.0%; Pred. No. 0.0058;
Matches 133; Conservative 0; Mismatches 114; Indels 4; Gaps 1;

Qy 3 ggaattgagggcggggtgctgcggtgtgctggggggcgccgagggcccgccggcg 62
Db 3478 GCGGTTCGCGCAGGGTGTGTGCGCGCGGTGCGGTCCGGCGCGCGCGCGCGCGCG 3419

Qy 63 gagcc-----ccggggagtcggcgctgctggagcgggtggtgctgagcgggcggtg 118
Db 3418 CGGCCAGGACCGGTCTGTCGGAGCGCGGTGAGCGCCACCGTGAAGCCGTCGCGGGCTGT 3359

Qy 119 gggccggcgggcgggagggcgccggggcgctgagcccccgggtacgaagatccggactccg 178
Db 3358 CGCGCGGCGAGTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3299

Qy 179 gcccgggactagagggcgctgcggcgctgagccactgtccaccacacacatgtgcagggc 238
Db 3298 GCTGCTTACCGAGGCGCGCTGCGCGCGCTGCTGAGCAGGACGAGGTTCGCGCGCGCG 3239
```



```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 816..14234
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14351..19945
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31232..36067
; NAME/KEY: CDS
; LOCATION: 36249..41774
US-08-804-227C-1

Query Match      8.7%; Score 50.6; DB 4; Length 5392;
Best Local Similarity 53.0%; Pred. No. 0.0058;
Matches 133; Conservative 0; Mismatches 114; Indels 4; Gaps 1;

QY 3 ggagttgaggggcgggtgtgcgctgtgacggcgccggcgccggcgccggcgccg 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3478 GGCGTTCGCGACGGTGTTGGCGCGGTGCGGTTCGCGCGCCGCCGCCGCCGAGCG 3419

QY 63 gaggc-----ccggggagtcggcgctgtgcgacgggttgcgacggcggttgggcccg 118
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3418 CGGCCAGGACCGGGTCTGCGGAGCCGCTGAGCGCCACCCTGAAGCGCTCGCGGGGCTGT 3359

QY 119 gggccggcgcggggagcgccggcgccgctgcagccccggtagacaagatccgactccg 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3358 CGCGGGCACTTCGGCGCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCT 3299

QY 179 gcccggaactagggcgctgcgcggtggagccactgtaccacgcacatggtggcgagcg 238
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3298 GCTGCTTACCGAGCGCGCTGCGGCCCTCTGTCGAGCAGGACGAGGTGCGGTGCGTCT 3239

QY 239 ccgtgagcggg 249
    | |||||
Db 3238 CGAGGCGCGGG 3228

RESULT 14
US-08-804-227C-1
; Sequence 1, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43280 base pairs
; TYPE: nucleic acid
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 03:49:39 ; Search time 4881.77 Seconds
(without alignments)
1609.094 Million cell updates/sec

Title: us-09-870-113-3
Perfect score: 582
Sequence: 1 atggagttgagggcgagg.....tgaacctgcggaaggtga 582

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	532.6	91.5	888	9	AL530804
2	450.6	77.4	817	10	BI103329
3	438.8	75.4	1131	10	BG295496
4	420.8	72.3	573	10	BI541703
5	409.8	70.4	697	10	BI854643
6	398.2	68.4	967	10	BI739710
7	387	66.5	645	10	BG964218
8	367.2	63.1	397	9	BE012485
9	354.6	60.9	574	10	BI185580
10	346.4	59.5	681	9	AV704087
11	345	59.3	598	10	BE913718
12	323.4	55.6	560	10	BI794616
13	320.2	55.0	480	10	BI045863
14	302.2	51.9	336	9	AW326482
15	298.4	51.3	546	10	BM488747
16	292	50.2	546	9	BE014800
17	290.8	50.0	900	10	BG325430

18	288.6	49.6	747	10	BI827322
19	282.8	48.6	501	9	AW211366
20	274.4	47.1	916	10	BI757180
21	255.8	44.0	622	10	BJ039504
22	244.6	42.0	722	10	BI876051
23	243.6	41.9	417	9	AA104365
24	240.8	41.4	855	10	BI412175
25	240.2	41.3	261	10	BI338246
26	231.4	39.8	701	10	BI831943
27	222.6	38.2	421	10	BI448355
28	207.4	35.6	578	9	BB654057
29	203.2	34.9	561	9	AI942584
30	196.6	33.8	684	9	AI133696
31	196.6	33.8	750	10	BI831750
32	194.6	33.4	729	10	BI460337
33	194	33.3	567	10	BJ035232
34	193.6	33.3	512	9	AW210205
35	191.6	32.9	453	10	BF916224
36	190.2	32.7	612	10	BJ012200
37	187.4	32.2	1295	11	AK015770
38	186.8	32.1	803	10	BG777404
39	184.6	31.7	217	9	AW346915
40	182.8	31.4	331	10	BI449976
41	181.6	31.2	1209	11	AK006155
42	177.8	30.5	666	9	BB644566
43	177.2	30.4	513	10	BI466991
44	174.2	29.9	866	10	BF797743
45	172	29.6	592	10	BM313240

ALIGNMENTS

RESULT 1

AL530804
LOCUS AL530804 LTI_NFL001_NBC4 Homo sapiens cdna clone CS0DD008YH17 5
DEFINITION prime, mRNA sequence.
ACCESSION AL530804
VERSION AL530804.1 GI:12794297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 888)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cdna libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DD008YH17"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cdna was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cdna was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

FEATURES
source

BASE COUNT

168 a 244 c 322 g 139 t 15 others

QY 528 cgggtgtggaacattacttcattgatgacagccatgaaccctgcggaag 577
||||| || ||||||||||||||||||||||||||||||| || |||||
Db 365 CGGGTGGT-ANACATTACTTCATGATCGACCCATGAATCCAGCAGAAG 413

RESULT 10
AV704087
LOCUS AV704087 681 bp mRNA linear EST 09-OCT-2000
DEFINITION AV704087 ADB Homo sapiens cDNA clone ADBAF05 5', mRNA sequence.
ACCESSION AV704087
VERSION AV704087.1 GI:10721407
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Peng,Y., Song,H., Huang,Q., Huang,C., Gu,Y., Yang,Y., Gao,G., Xiao
,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z.,
Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA ADB clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzq@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
Location/Qualifiers
1..681
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADBAFE05"
/clone_lib="ADB"
/tissue_type="Adrenal gland"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 149 a 209 g 117 t 1 others
ORIGIN

Query Match 59.5%; Score 346.4; DB 9; Length 681;
Best Local Similarity 99.7%; Pred. No. 8.1e-54;
Matches 347; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 116 gggggccggcgccgggagcgccgggctgcaggcccccgtacgacaagatccggact 175
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 GGGGGCCGGCGCGGGAGGCGCGGGCTGCAGGCCCGCGGTACGACAAATCCGGACT 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 176 ccggcccgactacaggcgctgcggctggagccactgtcaccccgacatggtggcag 235
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 CCGGCCCGGACTACAGAGCGCTGC CGCGGTGGAGCCACTGTCAACCGACATGTTGGCAG 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 236 gcgcctggcagggtacctgagacactcgctgatgccctacactgcgtccaagacc 295
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 GCGCCGAGCGAGGATCTGGAGCACTGGGTGATGTACCCCATCGACTCGGTCAAGACC 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 296 ggatcagagtacagcctgacccagctgcccgctatcgcaatgtgtggaggccctct 355
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 GGATCAGAGTCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCT 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 356 ggagattataaagacggagggcctatggagcccatagggggctgacgttcacagcaa 415
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 GGAGGATTATAAGAACGAGGGGCGCTATGGAGCCCATGAGGGGCTGAACGTCAAGCAA 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 416 caggcgaggccctgcaccacccctttatttgcctgtacgaaaaagt 463
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 CAGGCGCAGGGCGCTGCCACGCCCTTTATTTTGCTGTACGAAAAGT 348
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 11
BE913718
LOCUS BE913718 598 bp mRNA linear EST 29-SEP-2000
DEFINITION BE913718 601669072F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5',
mRNA sequence.
ACCESSION BE913718
VERSION BE913718.1 GI:10411618
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 598)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LRAM9146 row: o column: 17
High quality sequence stop: 593.
Location/Qualifiers
1..598
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969256"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 141 a 175 c 172 g 110 t
ORIGIN

Query Match 59.3%; Score 345; DB 10; Length 598;
Best Local Similarity 93.5%; Pred. No. 1.5e-53;
Matches 360; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 195 gctgcgctgagccactgtcacccacacatgtggcaggccgctggcaggatcct 254
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 GCTGCCGCTGAGGCCACTGTCAACCCACACATGTGGCGGCCGCTGGCAGGATCCT 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 255 ggagcactgcgtgatgtaccccatcactgcgtcaagaccggatgcagagtctacagcc 314
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 GGAGCATTTGCGTGATGTACCCGATCGACTGCGTCAAGACCCGGATGCAGAGCTACAGCC 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 315 tgaccacgtcccgctatcgcaatgtgtggaggccctctgaggatataagaacgga 374
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 TGACCCACCGCCCGCTATCGGAACGTGTGGAGCTCTCTGGAGAATCATGAGACAGA 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 375 ggccctatggagggccatgaggggtgaagtcacacacagcgagggcctgccca 434
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 GGGCCTGTGGAGGCCCATCGCGGGGCTGAACGTACACACACAGCGCGGGCCTGCCCA 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 435 gcgcctttatttgcctgtcacgaaagttaaaaaagacattgagtgaataccacc 494
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 CGCCCTCTATTGCTGCTACGAAAAGTTAAAAAGACATTGACTGACGTAATCCACC 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 495 tgggggcaatagcccatattgccaattggtcgggccgggtgtgtggaacattactcatga 554
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db	301	AGGGGCAATAGCCATTTCCTGCGGCGGATGTGCGGACATTACTTCATGA	360
Qy	555	tgagccatgaacctgcggaagc	579
Db	361	TGCAGCCATGAATCCAGCGAAGTC	385
RESULT 12			
BI794616			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
SOURCE			
BASE COUNT			

JOURNAL COMMENT

Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.
Location/Qualifiers

FEATURES

source
1. .546
/organism="Gallus gallus"
/strain="Commercial broiler and Ottawa Res. Centre
Strains 90 & 21"
/db_xref="taxon:9031"
/clone="pkm2n.pk008.12"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pkm2n)"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Breast, leg; Embryo(d19); post-hatch(1d,1,3,5,7,9
11 weeks); growth plate(1d,7d,14d post-hatch)"
/lab_host="E. coli EMDH10B"
/note="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"
90 a 183 c 197 g 70 t 6 others

BASE COUNT
ORIGIN

Query Match 51.3%; Score 298.4; DB 10; Length 546;
Best Local Similarity 77.0%; Pred. No. 4.6e-45;
Matches 375; Conservative 0; Mismatches 109; Indels 3; Gaps 1;
QY 91 ggggtggtcagcggcggtggcgggggcgggcgggggcgggcggtgcagg 150
DB 13 ggcggcgggcgggcgggcgggcgggcgggcgggcgggcgggcggtgcagg 72
QY 151 cccccggtacgacgaagatccggactccggcccggaactacagagcgctgcgggtggagcc 210
DB 73 ggcggcgggcgggcgggcgggcgggcgggcgggcgggcgggcggtgcagg 129
QY 211 actgtcacacgcacatggtggcgagcgccgtggcgaggtccctggagcactgcgtgatg 270
DB 130 gccgtgtccagcacatgctggcgggcgccgctggcgggcactatggagcactgcgtgatg 189
QY 271 taccctatcgactgcgtcaagaccggatgcagagtctacagctgacctgacctgcccgc 330
DB 190 taccccgctcgattgcgtcaagactcggatcgagcctgcggcgccgagcgccgcccgc 249
QY 331 tatcgcaatgttggaggccctctgaggattataagaacggagggcctatggaggccc 390
DB 250 tacccgaacgtgctggagggccctgtggcgcatctgtgcgtaccgagggcggtgtggagggcc 309
QY 391 atgagggggctgaacgtcacagcaacagcgcgagggcctgcccacgcccctttatttggc 450
DB 310 atggggggcctgaacatcacccgacccggcgggcgcccgcccgccctcttacttgcgc 369
QY 451 tgtctacaaaagttaaaaaagacattgagtgtatccaccctgggggcaatagccat 510
DB 370 tgcctacaaaagttaaaaaagacattgagtgtatccaccctgggggcaatagccat 429
QY 511 atgccaatggtgcggcggtgtgtggcaacattacttcatgatgcagccatgaaccct 570
DB 430 gtggcnnnccgtgcagcgggtgtgtgagcaacatttgcctccacgacgagggatgacccct 489
QY 571 gcggaag 577
DB 490 gcagag 496

Search completed: September 28, 2002, 06:22:39
Job time: 9180 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:53:07 ; Search time 5287.54 Seconds
(without alignments)
2742.691 Million cell updates/sec

Title: US-09-870-113-5
Perfect score: 693
Sequence: 1 atggagtggaggcgggg.....ttacctgtttaccactag 693

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10453268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	AF327402	83.3	1448	9	AF327402	Homo sapiens
2	AF327402	83.3	1448	17	AF327402	Homo sapiens
3	AF327402	83.3	1448	17	AF327402	Homo sapiens
4	AF327402	83.3	1448	17	AF327402	Homo sapiens
5	AF327402	83.3	1448	17	AF327402	Homo sapiens
6	AF327402	83.3	1448	17	AF327402	Homo sapiens
7	AF327402	83.3	1448	17	AF327402	Homo sapiens
8	AF327402	83.3	1448	17	AF327402	Homo sapiens
9	AF327402	83.3	1448	17	AF327402	Homo sapiens
10	AF327402	83.3	1448	17	AF327402	Homo sapiens
11	AF327402	83.3	1448	17	AF327402	Homo sapiens
12	AF327402	83.3	1448	17	AF327402	Homo sapiens
13	AF327402	83.3	1448	17	AF327402	Homo sapiens
14	AF327402	83.3	1448	17	AF327402	Homo sapiens
15	AF327402	83.3	1448	17	AF327402	Homo sapiens
16	AF327402	83.3	1448	17	AF327402	Homo sapiens
17	AF327402	83.3	1448	17	AF327402	Homo sapiens
18	AF327402	83.3	1448	17	AF327402	Homo sapiens
19	AF327402	83.3	1448	17	AF327402	Homo sapiens
20	AF327402	83.3	1448	17	AF327402	Homo sapiens
21	AF327402	83.3	1448	17	AF327402	Homo sapiens
22	AF327402	83.3	1448	17	AF327402	Homo sapiens
23	AF327402	83.3	1448	17	AF327402	Homo sapiens
24	AF327402	83.3	1448	17	AF327402	Homo sapiens
25	AF327402	83.3	1448	17	AF327402	Homo sapiens
26	AF327402	83.3	1448	17	AF327402	Homo sapiens
27	AF327402	83.3	1448	17	AF327402	Homo sapiens
28	AF327402	83.3	1448	17	AF327402	Homo sapiens
29	AF327402	83.3	1448	17	AF327402	Homo sapiens
30	AF327402	83.3	1448	17	AF327402	Homo sapiens
31	AF327402	83.3	1448	17	AF327402	Homo sapiens
32	AF327402	83.3	1448	17	AF327402	Homo sapiens
33	AF327402	83.3	1448	17	AF327402	Homo sapiens

ALIGNMENTS

RESULT 1

AF327402

LOCUS

DEFINITION

AF327402

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AF327402 Homo sapi
AJ303077 Homo sapi
AF327403 Homo sapi
AJ303078 Homo sapi
AX061229 Sequence
AL353719 Human DNA
AX071440 Sequence
AC096351 Rattus no
AC096351 Rattus no
AF288621 Mus muscu
AF223466 Homo sapi
AF361699 Mus muscu
AY032828 Homo sapi
BC015013 Homo sapi
AC099368 Rattus no
AC099368 Rattus no
AC108878 Mus muscu
AF216674 Homo sapi
AC051642 Homo sapi
AC022597 Homo sapi
AC306258 Sequence
Z66521 Caenorhabdi
AY060268 Drosophil
AP003518 Oryza sat
AF217402 Drosophil
AL396379 T7 end of
AC102327 Mus muscu
AC023381 Homo sapi
AC078884 Mus muscu
AC108558 Rattus no
AC095142 Rattus no
AC102213 Mus muscu
AC105592 Rattus no
AC034105 Homo sapi
AC012570 Homo sapi
AC021280 Homo sapi
AC105855 Rattus no
AC079411 Homo sapi
AC025066 Homo sapi
AC106354 Rattus no
M29384 Herpes simp
AC022552 Homo sapi
AC098461 Rattus no
AC097177 Rattus no
AC107039 Pan trogl
AC102552 Mus muscu

1448 bp mRNA linear PRI 02-MAY-2001
Homo sapiens putative mitochondrial solute carrier splice variant
mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.

AF327402
AF327402.1 GI:13926046

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1448)
Li.F.Y., Nikali.K., Grogan,J., Leibiger,I., Leibiger,B.,
Schweyen,R., Larsson,C. and Suomalainen,A.
Characterization of a novel human putative mitochondrial
transporter homologous to the yeast mitochondrial RNA splicing
proteins 3 and 4
FEBS Lett. 494 (1-2), 79-84 (2001)
21195335
11297739
2 (bases 1 to 1448)

AUTHORS	Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B., Schweyen, R., Larsson, C. and Suomalainen, A.
TITLE	Direct Submission
JOURNAL	Submitted (08-DEC-2000) Human Molecular Genetics, National Public Health Institute, Mannerheimintie 166, Helsinki 00300, Finland
FEATURES	Location/Qualifiers
source	1. .1448 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="10" /map="10q24" 30. .1124 /note="similar to mitochondrial RNA splicing protein 3/4; alternatively spliced" /codon_start=1 /product="putative mitochondrial solute carrier splice variant" /protein_id="AAK49519.1" /db_xref="GI:13926047" /translation="MELEGKAGCVAGGPAAGPGRSGESALLDGLQRCVGRGAGGG EAGACRPVRQDPDSDPEALFAGATVTHMVAVAGVILEHCVMYPIDCVKTRMOS LDTPAARYRNVAALRWLRIRITDEGLWRPMKRLNVKTAGPAHALYFACVEKLKLTLS DQVHGGSHIANAGCWRTLLHDAMNPAEVVKQOMYNSPYHRVTDCAVRVAMQ EGAGAFYRSYTTOLTWNVPQATHEMTYEFLOEHENPQRYNPSSHVLSCACAGAAA AATPLDVCKTLTNTQESIALNSHITGHGTGMASAFRTYVQGVGYATYFRGVQARYY QIPSTAIAKSVTEFFKYLITKKQEEWRACK"
CDS	322 a 408 c 435 g 283 t
BASE COUNT	
ORIGIN	

Query Match	83.3%	Score 577:	DB 9:	Length 1448:
Best Local Similarity	100.0%	Pred. No. 2.6e-88;		
Matches 577: Conservative	0;	Mismatches	0;	Indels
	0;	Gaps	0;	
QY	1	atgagttgagaggcggtgctgcgcggtgtgagcgggggccgcgcgcagagcccg	60	
DB	1			
DB	30	ATGGAGTTGAGGGCGGGGTCTGGCGGTGTGCGGGGGGGCGCGGACAGGCCCGGG	89	
QY	61	cgagacccccgggagtcggcgctgctggaacgggtggctcagcggggcgtgggcccgggg	120	
DB	1			
DB	90	CGGAGCCCGGGAGTCCGCGCTGCTGGACGGGTGGCTGCAGCGGGCGTGGGCGGGGG	149	
QY	121	gccgcgcgcgggagggccgggggacctgacagccccccggtacgacaagatccgggaactcggc	180	
DB	1			
DB	150	GCCGCGCGCGGGAGGCGCGGGGCGCTGCAGGCGCCCGGTACGACAAATCCGGACTCCGGC	209	
QY	181	ccggactacgagcgctgccgcggtggagccaactgtaccacgcacatggtggcagcgcc	240	
DB	1			
DB	210	CCGGACTACGAGCGCTGCCGGCTGGAGCCACTGTCTACCGCACATGGTGGCAGGCGCC	269	
QY	241	gtgcgagggatcctggagcaactgcgtgatatccccaatgactgcgtcaagacccgatg	300	
DB	1			
DB	270	GTGGCAGGGATCTGGAGCACTGCGTGATGTACCCCATCGACTGCGTCAAGACCCGATG	329	
QY	301	cagagttacagctcgtaccacagctgccgcgtatctcgaaatgtgtggaggccctctggag	360	
DB	1			
DB	330	CAGAGCTACAGGCTGACCCAGCTGCCCGCTATCGCAATGTGTGGAGGGCCCTCTGGAGG	389	
QY	361	attataagaacgagggccctatggagggcccatggaggggctgaacgtcacagacaacaggc	420	
DB	1			
DB	390	ATTATAAGAACGGAGGGCCCTATGGAGGCCCATGAGGGGGCTGAACGTACAGCAACAGGC	449	
QY	421	gcagggcctgccacgcgccttattttgctgctacgaaaagttaaaaaaacattgaagt	480	
DB	1			
DB	450	GCAGGGCTGCCACGCGCCCTTATTTTGCTGCTACGAAAGTTAAAAAAGACATTGAGT	509	
QY	481	gatgtaatccaccctggggcaatgacatattgcaaatggtgcgcgcggtgtgtggca	540	
DB	1			
DB	510	GATGTAATCCACCCTGGGGCAATAGCCATATTTGCCAATGGTGGCGCGGGTGTGGCA	569	
QY	541	acattactteatgatgcagccatgaacccctgcggaag	577	
DB	1			
DB	570	ACATTACTTTCATGATCAGCCATGAACCCCTTCGCGAAG	606	

RESULT 2

HSA303077 standard; RNA; HUM; 1448 BP.

XX AC AJ303077;

XX SV AJ303077.1

SV XX

DT DT 02-FEB-2001 (Rel. 66, Created)

TT DT 26-APR-2001 (Rel. 67, Last updated, Version 2)

XX XX

DE DE Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMR33/4

DE gene), 1448 bp cDNA splice variant

XX XX

KW HMR33/4 gene; mitochondrial RNA splicing protein 3/4.

XX XX

OS Homo sapiens (human)

OS OS

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC OC

XX Eutheria; Primates; Catarrhini; Homnidae; Homo.

XX XX

[1] 1-1448

RN RP RA Nikali K.;

RT ;

RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.

RL RL Nikali K., Human Molecular Genetics, National Public Health Institute,

RL RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.

XX XX

[2]

RN Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,

RN RA Larsson C., Suomalainen A.;

RT "Characterization of a novel mitochondrial transporter

RT RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4".

FEBS Lett. 494:79-84(2001).

XX XX

Key Location/Qualifiers

FH FH

FT source 1. .1448

FT FT /db_xref="taxon:9606"

FT FT /organism="Homo sapiens"

FT FT 30. .1124

CDS

FT FT /gene="HMR33/4"

FT FT /product="mitochondrial RNA splicing protein 3/4"

FT FT /function="putative mitochondrial solute carrier"

FT FT /protein_id="CAC27996.1"

FT FT /translation="MELEGRGAGVGACGPGSPGESALLDGHQGRGVGRGAGGE

FT FT AGACRPVRDDPSGGPEALPAGAVITTHMAGVAGILEHCVMYIPDCVKRMOSLO

FT FT PDPAARYRNVLLEALRWIRTEGLWRPMRGLNTATGAGPAHALYFACYEKLLKLSLDI

FT FT HPGNSHIANGAAGCAVTLLHDAMNPAEVVKQMOMNSPHRVYTDVCVRAVQNEGAG

FT FT AFYSRYTTLTMNVPPFOA IHMTVEFLQEHFNPRRY NPSSHVLSGCAGAAAAATT

FT FT LDVCKTLLNQESLANSHITGHTMASAFRTVYQVGGVAYFRGVQRVIYQIPSTA

FT FT IAWSVYEFFKYLIHKROEWRACK"

XX XX

Sequence 1448 BP; 322 A; 408 C; 435 G; 283 T; 0 other;

	Query Match	83.3%	Score 577;	DB 17;	Length 1448;
	Best Local Similarity	100.0%;	Pred. No. 2.6e-88;		
	Matches 577;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	atggaattgagggcggggtgctgacggtatgacggggggggcgcgcgagggcccggg	60		
DB	30	ATGGATTGGAGGGCGGGGTCTGGCGGTGTGGGGGGGGCGCGGAGGCGCGGG	89		
QY	61	cggagcccccgggagtcgagcgctgctggacgggtggctgcagcggggcgctgggcccgggg	120		
DB	90	CGAGCCCCGGGAGCTCGCGCTCTGGACGGGTGGCTGTGACGGGGCGTGGCCCGGGG	149		
QY	121	gcggcgcgcgggagggcgggcggtcgaggcccccggttaoagacaagatccggactccgc	180		
DB	150	GCGGCGGCGGGAGCGCGGGGCTGTGAGGCCCGCGGTACACAAGATCCGGACTCCGC	209		

RL Nikali K., Human Molecular Genetics, National Public Health Institute,
 RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.
 XX [3]
 RN Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
 RA Larsson C., Suomalainen A.;
 RA "Characterization of a novel human putative mitochondrial transporter
 RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";
 RL FEBS Lett. 494:79-84(2001).
 XX
 FH Key Location/Qualifiers
 FH source 1..1889
 FT /db_xref="taxon:9606"
 FT /organism="Homo sapiens"
 FT CDS 1032..1565
 FT /gene="HMR3/4"
 FT /product="mitochondrial RNA splicing protein 3/4"
 FT /function="putative mitochondrial solute carrier"
 FT /protein_id="CAC27997.1"
 FT /translation="MNPAEVVKQRMQVNSPYHRVTDVRAVWNEGAGAFYRSYTTQL
 TMVPPQAIHFMTYEFLOEHFNQRRYNPSHVLSCAGACAVAAATPLDCKTLLNT
 QESLAINSHITGHITGMASAFRTYVGGVTAFRGVQARVIQIPSTAIAMSVIEFFK
 YLITKQEWRAK"
 FT
 XX
 SQ Sequence 1889 BP: 404 A; 508 C; 544 G; 433 T; 0 other;

Query Match 75.1%; Score 520.4; DB 17; Length 1889;
 Best Local Similarity 99.8%; Pred. No. 9.4e-79;
 Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atgaggttgagggcggggtgctggcggtggtggcgggggcgcgcgagggcccg 60
 Db 9 ATGAGGTGAGGGCGGGGTGCTGGCGGTGTCGGGGGGGCGAGCGGCGCGG 68
 Qy 61 cggagccccgggagtcggtgctggaaggggtggtgagcgggggtggtggcg 120
 Db 69 CGGAGCCCCGGGAGTGGCGTCTGTGACGGGTGGCTGCAGCGGGGCGTGG 128
 Qy 121 gccggcggggagggcgggcggtgctgagcccccgcggtacgacagatccgg 180
 Db 129 GCCGGCGGGGAGGCGGGGGCTGCAGGCCCCCGGTACGACAAATCCGGCT 188
 Qy 181 ccggactacagggcggtgctggcggtgagccactgtcaccagcaatggtggc 240
 Db 189 CCGGACTACGAGGGCGTCCGGCTGGAGCCACTGTACACGACACATGTTGG 248
 Qy 241 gtggcagggatcctggagcactgctgtagtaccatcgactgctgcaagaccgg 300
 Db 249 GTGGCAGGGATCCTGGAGCACTGGGTGATGTACCCCATCGACTCGCTCA 308
 Qy 301 cagagttacagcctgacccagctgcccgtatgcaaatgtgttgaggcccttg 360
 Db 309 CAGAGTCTACAGCCTGACCCAGCTGCCCGCTATGCGAATGTGTTGGAGGC 368
 Qy 361 attataaagcagggcgcttatgagcccatgaggggctgacgtcacgacaagc 420
 Db 369 ATTATAGACGAGGCGCTATGAGGCCCCATGAGGGGGCTGACGACGACAG 428
 Qy 421 gcaggcgctgccacgcctctttatttgcctgctacgaaaagttaaaagacatt 480
 Db 429 GCAGGGCTGCCACGCGCTTTATTTGCTGCTGCTACGAAAAGTTAAAAAG 488
 Qy 481 gatgtaataccactgggggcaatagcattatgccaatggt 522
 Db 489 GATGTAATCCACCTGGGGGCAATAGCCATATTGCCCAATGTT 530

RESULT 5
 AX061229 AX061229 1322 bp DNA linear PAT 22-JAN-2001
 LOCUS
 DEFINITION Sequence 76 from Patent WO0078953.

AX061229
 VERSION AX061229.1 GI:12406365
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1322)
 AUTHORS Lal,P., Yang,J., Yue,H., Hillman,J.L., Tang,Y.T., Bandman,O.,
 Burford,N., Baughn,M.R., Azimzai,Y., Lu,D.A., Au-Young,J. and
 Patterson,C.
 TITLE Human transport proteins
 JOURNAL Patent: WO 0078953-A 76 28-DEC-2000;
 Incyte Genomics, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..1322
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 328 a 376 c 351 g 267 t
 ORIGIN

Query Match 65.6%; Score 454.4; DB 6; Length 1322;
 Best Local Similarity 98.7%; Pred. No. 1.4e-67;
 Matches 458; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 114 ccggggggcgcgcggggagggcggtgctggcggtgagccactgtcaccaacgcatggtgc 173
 Db 2 CGGGGGCGCGCGGGGAGGCGGGGCTGCAGGCCCGCCCGGTAGCAAGATCCGGA 61
 Qy 174 ctccggggcggtacgagggcggtgctggcggtgagccactgtcaccaacgcatggtgc 233
 Db 62 CTCGGGCGCGGACTACGAGCGCTGCGGGCTGGAGCCACTGTCAACACGACATGGTGG 121
 Qy 234 aggcgcgtgaggggactcctgagcactgcgtgagtaccccatgagctgcacagac 293
 Db 122 AGCGCGCGTGGCAGGGATCCTGGAGCACTGCGTGTATGCCCATCGCTACGCTCAAG 181
 Qy 294 ccggatcgagctgtacagctgacccagctgccgctatcgcaatgtgtgagggccct 353
 Db 182 CCGGATGCAGAGTCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCT 241
 Qy 354 ctggagggattataagaacgagggcgcttatggagggccatgagggggctgcaacgtcac 413
 Db 242 CTGGAGGATTATAAGAACGAGGAGCGCTTATGGAGGCCCATGAGGGGGCTGAACGTCAC 301
 Qy 414 aacaggcgagggcgctgcccacgccccctttatttgcctgctagaaaaagattaaaaa 473
 Db 302 AACAGGCGCAGGGCGCTGCCACGCGCTTTATTTGCTGTCTACGAAAAGTTAAAAAG 361
 Qy 474 attgagtgatgtaatccacccctggggcaatagccatattgccaatggtgcggcggtg 533
 Db 362 ATTGAGTGTATGTAATCCACCTTGGGGCAATAGCCATATTGCCCAATGTCGCGCGGG 421
 Qy 534 tgtggcaacattacttcattgatgcagccatgaacccctgcggaag 577
 Db 422 TGTGGCAACATTACTTTCATCATGCAGCCATGAACCCCTGCGGAAG 465

RESULT 6
 AL353719/c
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete
 sequence.
 ACCESSION AL353719 AC007643
 VERSION AL353719.10 GI:15787725
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 123160)
 AUTHORS Ramsay,H.

REFERENCE AUTHORS

Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaraturunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brleva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.K., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havliak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogulu,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
32851: contig of 32851 bp in length
32852: gap of unknown length
32951: contig of 16033 bp in length
48984: gap of unknown length
48985: gap of unknown length
49084: gap of unknown length
49085: contig of 14731 bp in length
63915: gap of unknown length
63916: contig of 12090 bp in length
76005: gap of unknown length
76105: contig of 12736 bp in length
88841: gap of unknown length
88842: gap of unknown length
88942: contig of 11650 bp in length
100591: gap of unknown length
100691: contig of 11011 bp in length
100692: gap of unknown length
111703: contig of 10862 bp in length
111803: gap of unknown length
122665: gap of unknown length
122765: contig of 6947 bp in length
129711: gap of unknown length
129712: contig of 7436 bp in length
137248: gap of unknown length
137348: contig of 7329 bp in length
144676: gap of unknown length
144677: contig of 7348 bp in length
152124: gap of unknown length
152125: contig of 6654 bp in length
152225: gap of unknown length
158978: contig of 4950 bp in length
158979: gap of unknown length
163929: contig of 5325 bp in length
164029: gap of unknown length
169354: contig of 6137 bp in length
169454: gap of unknown length
175590: contig of 4785 bp in length
175691: gap of unknown length
180476: contig of 4082 bp in length
180576: gap of unknown length
184658: contig of 5575 bp in length
184758: gap of unknown length
190333: contig of 3702 bp in length
190433: gap of unknown length
194135: contig of 2358 bp in length
194235: gap of unknown length
196593: contig of 3838 bp in length
200530: gap of unknown length
200630: contig of 2179 bp in length
202810: gap of unknown length
202811: contig of 2125 bp in length
205034: gap of unknown length
205134: contig of 2042 bp in length
207176: gap of unknown length
207277: contig of 1569 bp in length
208845: gap of unknown length
208945: contig of 2326 bp in length
211271: gap of unknown length
211371: contig of 1931 bp in length
213302: gap of unknown length
213403: contig of 2082 bp in length
215485: gap of unknown length
215485: contig of 1385 bp in length
215855: gap of unknown length
216970: contig of 1289 bp in length
217070: gap of unknown length
218359: contig of 1428 bp in length
219887: gap of unknown length
219887: contig of 1076 bp in length.
Location/Qualifiers
1. .221062
/organism="Rattus norvegicus"

FEATURES
source

```

/db_xref="taxon:10116"
/chromosome="Rf1"
/clon="CH230-24M6"
BASE COUNT 58389 a 50577 c 49732 g 58877 t 3487 others
ORIGIN

Query Match 36.2%; Score 251; DB 2: Length 221062;
Best Local Similarity 91.4%; Pred. No. 2.le-33;
Matches 266; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 atgagttgagggcggtgtgctggtgagcgtgtgagcggggggcgcgagggcccggg 60
|||||
Db 72865 ATGAGTTGAGGGCGGAGTGCGAGCGGCGTGGCGGAGGACGAGCTGTGGCCCGGG 72924
|||||

QY 61 cggagccccgggagtcggtgagcgtgctgagcgttggtgcagcggggcggtggcgccgggg 120
|||||
Db 72925 CGGAGCCCCGGGAGTGGCGCTGCTGGACGGGTGGCTGCAGCGGGCGTGGCGCGGGG 72984
|||||

QY 121 gccggcgccggggagcgccgggctgcagcccccgcgtgacgacgaagatccggactccggc 180
|||||
Db 72985 GCGCGCGCGGGAGCGGGGGCCCTATCAGACCCCTGTACGGCTGGATCCGAGTCCGGT 73044
|||||

QY 181 ccggactacagggcgtgcggctggagccactgtcaccacgcacatggtggcagggccc 240
|||||
Db 73045 CCGGAATACGAAGCGCTGCCGGCTGGAGCCACTGTCTACCGACATGCTGGCGGGGCC 73104
|||||

QY 241 gtgcagggatcctggagcactgcgtgatgtacccctcagctcgtcgaag 291
|||||
Db 73105 GTGCGAGGATCTGGAGCATGTCGCTGATGTACCGATCGACTCGGTCAG 73155
|||||

RESULT 9
AC096351/c
LOCUS
DEFINITION Rattus norvegicus chromosome Rf1 clone CH230-24M6, WORKING DRAFT
SEQUENCE, 33 unordered pieces.
AC096351
VERSION AC096351.2 GI:17944054
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimge,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Deen,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Haves,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,D.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,

```

```

Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Teifrod,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
* 32851: contig of 32851 bp in length
* 32951: gap of unknown length
* 32952: contig of 16033 bp in length
* 48985: gap of unknown length
* 49085: contig of 14731 bp in length
* 63815: gap of unknown length
* 63916: contig of 12090 bp in length
* 76006: gap of unknown length
* 76106: contig of 12736 bp in length
* 88841: gap of unknown length
* 88941: contig of 11650 bp in length
* 100592: gap of unknown length
* 100692: contig of 11011 bp in length
* 111703: gap of unknown length
* 111802: contig of 10862 bp in length
* 122664: gap of unknown length
* 122764: contig of 6947 bp in length
* 129715: gap of unknown length
* 129711: contig of 7436 bp in length
* 129812: gap of unknown length
* 137248: contig of 7329 bp in length
* 137348: gap of unknown length
* 144676: contig of 7348 bp in length
* 144777: gap of unknown length
* 152124: contig of 7348 bp in length
* 152125: gap of unknown length

```



```
Db 464 GCGCAGCCCATGTAATTTGCTGCTATGAAACATGAAAGGACATTAATAAGACGTTTTC 523
QY 490 caocctgggggcaatagccatattgccaatgtgagcgccgggtgtgtgcaacattactt 549
Db 524 AGCCACCAAGGAAGAACAGCCATAGCTAATGGGTAGCTGGGAGTATGCCACCCCTACTC 583
QY 550 catgatgagccatgaacccctcggaag 577
Db 584 CACGATGCAGTAATGAATCCAGAGAAG 611

RESULT 11
LOCUS AF223466 1429 bp mRNA linear PRI 17-APR-2000
DEFINITION Homo sapiens HT015 protein (HT015) mRNA, complete cds.
ACCESSION AF223466
VERSION AF223466.1 GI:7578782
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1429)
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
TITLE A novel gene expressed in human hypothalamus
JOURNAL Unpublished
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-2000) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, China
FEATURES
source
1..1429
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
1..1429
/gene="HT015"
45..1088
/gene="HT015"
/codon_start=1
/product="Hr015 protein"
/protein_id="AAF64141.1"
/db_xref="GI:7578783"
/translation="MELRSGSVGQAVARRMDGSDRGCGKDATGSDYENLPTSAS
VSTHTAGAMAGILEHSMYVDSVKTRQSLSDPPKQAYTSIYGALKIRITEGFWR
PLRGVNVIMAGPAHAMAFYACIENMKRTLNDVFRHQNSHLANGIAGMATLLHDV
MNPAEVVKRLOMTYNQSHSAISCIRTVWTEGLGAFYRSTLTMTNIPFQSIHFIT
YEFLQEQVNPRTYNQSHSIISGLAGLAAAATPLDVKCTLNTQENVALSANIS
GRLSGMANAFRTVQLNGLPATSKASRRVSTRCPPLPFLGLSMSSSTFSPASWKI
ELHTKGRDHRIFS"
BASE COUNT 326 a 426 c 402 g 275 t
ORIGIN

Query Match 28.7%; Score 198.8; DB 9; Length 1429;
Best Local Similarity 62.7%; Pred. No. 2.4e-24;
Matches 325; Conservative 0; Mismatches 192; Indels 1; Gaps 1;

QY 61 cggagcccccgggagtcgcgctgtgtgacgggtgtgagcgccggcggtggcgccgggg 120
Db 23 CTGCGCCCGCGGAGTGGCGGATGAGCTGCGCAGCGGCGGTGGCGAGCCAGCGCGT 82
QY 121 qccgagcggcgaggcgccgtgcaggcccccggtacgacagatccgagctccggc 180
Db 83 GCGCGGAGGATGGATGGGACAGCGGAGATGGCGCGCGGCGCAAGGACGCCCGGGTC 142
QY 181 -ccggactacgagcgctgcggcggtggagccactgtccaccgacacatggtgagcgcc 239
Db 143 GGAGGACTACGAGAACCTGCCGACTAGCGCTCCGTGTCCACCCACATGACAGCAGGAC 202
QY 240 cgtggcagggatcctggagcactgcgtgatgtacccccatcgactgcgtcaagaccggat 299
```

```
Db 203 GATGGCCGGGATCCTGGAGCACTCGCTATGACCCGGTGGACTCGGTGAAGACACAGAAAT 262
QY 300 gcagagctctacagcctgacccagctgccgctatcgcaatggtgtgagccctctggag 359
Db 263 GCAGAGTTTGAGTCCAGATCCCAAGCCCAAGTACACAAAGTATCTACGGAGCCCTCAAGAA 322
QY 360 gattataagaacggaggccctatggagggcccatgaggggctgaacgtcacagcaacagg 419
Db 323 AATCATGCGGACCGAAGGCTTCTGGAGGCCCTTGGAGCGCTCAACGCTCATGATCATGGG 382
QY 420 cgcagggcctgccacgcccctttattttcctgctactacgaaaagttaaaaaagacattgag 479
Db 383 TGCAGGGCCGGCCCATGCCCATGTATTTTCCTGCTATGATAAATGATAAAGGACTTTAAA 442
QY 480 tgatgtaatccaccctcggggccaatagccatattgccaatggtgcggccggtgtgtgac 539
Db 443 TGACGTTTTCACCAAGGAAACACACCACCTAGGCAACGGGATAGCTGGGAGTATGGC 502
QY 540 aacattactcatgatgcagccatgaacccctcggaag 577
Db 503 CACCCTGCTCCAGGATGCGGTAATGAATCCAGCAGAAG 540

RESULT 12
LOCUS AF361699
DEFINITION Mus musculus mitochondrial carrier-like protein mRNA, complete cds;
nuclear gene for mitochondrial product.
ACCESSION AF361699
VERSION AF361699.1 GI:16755527
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 5869)
AUTHORS Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and
She,J.
TITLE A novel mouse mitochondrial carrier protein gene is up-regulated
from young to adult NOD mice
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5869)
AUTHORS Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and
She,J.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2001) Department of Pathology, Immunology and
Laboratory Sciences, University of Florida, 1600 SW Archer Road,
Room D6-15, Gainesville, FL 32610, USA
FEATURES
Location/Qualifiers
1..5869
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="14"
/map="between D14Mit124 and D14Mit236"
116..664
/note="MCLP"
/codon_start=1
/evidence="not_experimental"
/product="mitochondrial carrier-like protein"
/protein_id="AAL27990.1"
/db_xref="GI:16755528"
/translation="MELRGGVGNQAGRRMDGDCRDGCGSKDAGSDYENLPTSAS
VSTHTAGAMAGILEHSMYVDSVKTRQSLSDPPKQAYTSIYGALKIRITEGFWR
PLRGVNVIMAGPAHAMAFYACIENMKRTLNDVFRHQNSHLANGIAGMATLLHDV
YEFLQEQVNPRTYNQSHSIISGLAGLAAAATPLDVKCTLNTQENVALSANIS
GRLSGMANAFRTVQLNGLPATSKASRRVSTRCPPLPFLGLSMSSSTFSPASWKI
ELHTKGRDHRIFS"
BASE COUNT 1387 a 1427 c 1390 g 1665 t
ORIGIN

Query Match 25.7%; Score 177.8; DB 10; Length 5869;
Best Local Similarity 63.6%; Pred. No. 7.3e-21;
```


Matches 288; Conservative 0; Mismatches 162; Indels 3; Gaps 1;

```
Qy 73 gagtcgagcgtctgagacgggtggtgtgacgagggcggtggtgagggggggcgagggcggg 132
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 104 GCGGGGCGCGGATGAGGCTCCGCGGTGCGGCGGTGGGGAACACAGGCTCGGGCGGAGG 163
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 133 gaggcgcgggctcagggcccccggtacgacaaatcgactcgactcgcc---ggactac 189
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 164 ATGGATGGGACTCGCGATGGCGGTGCGGCACCAAGGACCGCGGTGCGAGGACTAC 223
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 190 gaggcgctgcccgtggagccactgtcaccacgcacatgtggcgagggcgccgtgacagg 249
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 224 GAGAACCTACCCAGCACGCGCTCGGTGTCCACCCACATGACACGGGAGCGATGCCCGG 283
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 250 atccctggagcactcggtatgtacccctcgactcgctcaagagaccggatgcagagtcta 309
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 284 ATTCGGAGCACTCCATGATGATACCGGTGGACTCCGTGAAGACACGGATGACAGAGTTG 343
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 310 cagctgacccagctgcccgtctatcgcaatgtgttggaggccctctgaggattataaga 369
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 344 ATCCAGATCCCAAGCCGGTATACAGCATCTATGCGGCCCTCAAGAGGATCATGCAC 403
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 370 acggaggccctatgagggccatgaggggctgaaactgcacagcaacagggcgagggcct 429
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 404 ACTGAAGGCTTCGGAGGCGCCCTGCGGGCCCTGGAACGTGATGATGGGTGACGGGCC 463
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 430 gccacgccttatttgcctgtacgaaaagttaaagacattgagtgatgaatc 489
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 464 GCGCAGCGCATGATTTTGGCTGTATGAAACATGAAAGGACTTTAAATGACGTTTTC 523
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 490 caccctgggggcaatagccatattcccaatgt 522
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 524 AGCCACCAAGGAACAGCCATCTAGCTAATGT 556
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 13

```
AY032628          3415 bp  mRNA  linear  PRI 23-APR-2001
LOCUS              Homo sapiens mitochondria solute carrier protein (MSCP) mRNA,
DEFINITION          complete cds, alternatively spliced.
ACCESSION            AY032628
VERSION              AY032628.1  GI:13774108
KEYWORDS              human.
SOURCE                human.
ORGANISM              Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE            1. (bases 1 to 3415)
AUTHORS              Li,Q.-Z., Eckenrode,S., Ruan,Q.-G., Wang,C.-Y., Shi,J.-D.,
                     McIndoe,R.A. and She,J.-X.
TITLE                Molecular cloning of a novel mitochondria solute carrier protein
                     (MSCP) gene from mouse and human and its down-regulation in mouse
                     spleen during the maturation of the immune system
JOURNAL              Unpublished
REFERENCE            2. (bases 1 to 3415)
AUTHORS              Li,Q.-Z., Eckenrode,S., Ruan,Q.-G., Wang,C.-Y., Shi,J.-D.,
                     McIndoe,R.A. and She,J.-X.
TITLE                Direct Submission
JOURNAL              Submitted (15-APR-2001) Department of Pathology, Immunology and
                     Laboratory Sciences, University of Florida, 1600 SW Archer Road,
                     Room D6-15, Gainesville, Florida 32610, USA
FEATURES
   source             1. 3415
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p21"
   gene               1. 3415
                     /gene="MSCP"
   CDS                45..512
                     /gene="MSCP"
                     /note="alternatively spliced"
                     /codon_start=1
```

```
/product="mitochondria solute carrier protein"
/protein_id="AAK38154.1"
/db_xref="GI:13774109"
/translation="MEIURSGVGSQAVARMGDSRGGGKDATGSDYENLPTSAS
VSTHVMAGMAGLLEHSMYPSVKTRMOSLSPDKAQTSTYIYALKIKMRTGFWR
PLRGVNMIMGAPAHAMFYACVEMFRTLNDVFHOGNSHLANGILKAFVWS"
BASE COUNT          799 a 851 c 846 g 919 t
ORIGIN
```

Query Match 25.1%; Score 174.2; DB 9; Length 3415;
Best Local Similarity 62.4%; Pred. No. 3.1e-20;
Matches 289; Conservative 0; Mismatches 173; Indels 1; Gaps 1;

```
Qy 61 cggagccccgggagtcggcgctgctggacgggtggtgtgacgagggcggtgagggcgggg 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 23 CTGGCCCCCGCGAGCTGGCGGATGAGCTGCGACGGGAGCGGTGGCGACGCCAGCGCGT 82
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 121 gccggcgcggggagggcgccgtgcaggcccccggtcacgacagatccggactccggc 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 83 GCGCGGAGGATGGATGGGACAGCCGAGATGGCGCGCGGCGCAAGAGCGCCACCGGTC 142
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 181 -ccggactacgagcgctgcccgggtggagccactgtgaccacgcacatggtggcaggcgc 239
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 143 GGAGGACTTACGAGAACCTTGGCGAGTAGCGCTCGGTGCCACCCACATGACAGCAGGAGC 202
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 240 cgtggcagggaactcgagcactcggtatgtacccctcgactcgctcaagaccggat 299
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 203 GATGCGCGGATCTGGAGCAGCTCGGTGATGTACCGGTGGACTCGGTGGAAGACACGAAT 262
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 300 gcagagctacagctgacccagctgcccgcctatcgcaatgtgttggaggccctctggag 359
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 263 GCAGATTTGAGTCCAGATCCCAAGCCAGTACACACAGTATCTACGAGGCCCTCAAGAA 322
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 360 gattataagacgagggcctatggagggccatgaggggctgaacgtcacagcaacagg 419
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 323 AATCATGCGGACCGAAGCTTCTGGAGGCCCTTGGAGGCCGTCAACGTCATGATCATGGG 382
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 420 cgcaggcctcccacgccccttatttgcctgtacgaaaagttaaagacattgag 479
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 TGCAGGGCCAGCCCATGCTGATTTTGGCTGTATGAAACATGAAAGGACTTTAAA 442
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 480 tgatgtaatccaccctggggcaatagccatattgccaatgt 522
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 443 TGACGTTTCCACCACCAAGGAACAGCCACCTAGCCAAACGGT 485
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 14

```
BC015013          858 bp  mRNA  linear  PRI 04-OCT-2001
LOCUS              Homo sapiens, clone MGC:8764 IMAGE:3910567, mRNA, complete cds.
DEFINITION          BC015013
ACCESSION            BC015013
VERSION              BC015013.1  GI:15929109
KEYWORDS              MGC.
SOURCE                human.
ORGANISM              Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE            1. (bases 1 to 858)
AUTHORS              Strausberg,R.
TITLE                Direct Submission
JOURNAL              Submitted (01-OCT-2001) National Institutes of Health, Mammalian
                     Gene Collection (MGC), Cancer Genomics Office, National Cancer
                     Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
                     USA
REMARK              NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT              Contact: MGC help desk
                     Email: qcabs-r@mail.nih.gov
                     Tissue Procurement: ATCC
                     cDNA Library Preparation: Life Technologies, Inc.
                     CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                     DNA Sequencing by: Genome Sequence Centre,
                     BC Cancer Agency, Vancouver, BC, Canada
```



```

info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 15 Row: h Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7706149.
FEATURES
Location/Qualifiers
1..858
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:8764 IMAGE:3910567"
/tissue_type="uterus, leiomyosarcoma"
/clone_lib="NIH MGC_71"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
26..493
/codon_start=1
/product="Unknown (protein for MGC:8764)"
/protein_id="AAH15013.1"
/db_xref="GI:15929110"
/translation="MELRSGVSGQAVARRMDGSDRGGGKDATGSDYENLPTSA
VSTHGMVIMAGALEHSPVDSVKTRMQSLSPDKPAQYTSYIGALKIMRTGFWR
PLRGVNMIMAGFAHAMFYACENMKRTLNDVPHOGNHSLANGILKAFVMS"
BASE COUNT 215 a 220 c 231 g 192 t
ORIGIN

Query Match 24.8%; Score 172; DB 9; Length 858;
Best Local Similarity 69.1%; Pred. No. 8.5e-20;
Matches 235; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 183 ggaactacgagcgctgcggctggagccactgtcaccagcaatggtggcaggcgccgt 242
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 GGACTACGAGAACCTGCGGACTAGCGCTCCGTGTCACCCACATGACAGCAGGACCGAT 186
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 243 ggcaggatcctggagcactgcgtgatgtacccctgcactcgtcaagaccggatgca 302
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 GCGCGGATCTGGAGCACTCGGTATGACCGGTGACTCGGTGACACACGACGATGCA 246
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 303 gagttacagcctgaccagctgcccgcgtatcgcgaatgtgttggaggccctctggagat 362
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 GAGTTTGAGTCCAGATCCCAAGCCAGTACACAAGTATCTACGGAGCCCTCAAGAAAAAT 306
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 363 tataaagacgagcgccctatgagggccatgaggggctgaacgtcacagcaaggcgc 422
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 CATGCGGACCGAAGGCTCTTGGAGGCGCTTGGAGCGGTCAACGTCATGATCATGGGTGC 366
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 423 agggcctgcccacgccttttttgcctgcacgaaagttaaaaaagattgattgagta 482
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 AGGCGCGGCCATGCCATGATTTTTCCTGCTATGAARACATGAAGAGCATTAAATGA 426
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 483 tgtaataccacctggggcaatagccattatgccaatggt 522
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 CGTTTTCACCAACCAAGAAACAGCCACCTAGCAGCGT 466
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
AC099368
LOCUS
DEFINITION Rattus norvegicus clone CH230-84G11, *** SEQUENCING IN PROGRESS
***, 71 unordered pieces.
ACCESSION AC099368
VERSION AC099368.2 GI:17973175
KEYWORDS HTG; HTGS_PHASE1.

```

SOURCE ORGANISM

Norway rat.

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

1 (bases 1 to 163205)

REFERENCE AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbarella,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burrell,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
Hollins,B., Homsif., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherez,S., Scott,G., Shen,H., Shoohtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission

2 (bases 1 to 163205)

REFERENCE

Worley,K.C.

Direct Submission

TITLE

Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16901901.

COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHJ
Center clone name: CH230-84G11
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 127058 bases at least Q40
Consensus quality: 138264 bases at least Q30
Consensus quality: 146548 bases at least Q20
Estimated insert size: 120308; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 08:05:43 ; Search time 664.31 Seconds
(without alignments)
1791.064 Million cell updates/sec

Title: US-09-870-113-5
Perfect score: 693
Sequence: 1 atggagtggaggcgggg.....ttacctgtttaccactag 693

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length DB ID	Description		
1	469.4	67.7	1336 21 AAC90457	Human uncoupling p		
2	458	66.1	1418 21 AAC74843	Human ORFX ORF398		
3	454.4	65.6	1322 22 AAF27733	Human transport pr		
C	4	440	63.5 1294 22 AAI60661	Human polynucleoti		
	5	401	57.9 2502 22 AAS03906	Human secreted pro		
	6	398.6	57.5 1316 22 AAS03906	Human polynucleoti		
	7	273	39.4 401 22 AAF66156	Novel human polynu		
8	233	33.6	27960 22 AAK69779	Human immune/haema		
9	233	33.6	27960 22 AAK73320	Human immune/haema		

10	200.4	28.9	1902	22	AAK521172	Human polynucleoti
11	196.6	28.4	1305	22	AAF59920	Human mitochondria
12	187.4	27.0	2562	22	AAH90057	Human bone marrow
13	181	26.1	339	22	AAK58672	Human immune/haema
14	174.2	25.1	1890	22	AAK94897	Human full-length
15	173.6	25.1	842	22	AAK92357	Human cDNA 5'-end
16	173.6	25.1	842	22	AAK93895	Human cDNA clone r
17	173.6	25.1	1814	22	AAK94871	Human full-length
18	172.6	24.9	1468	21	AAK971175	Human ORFX ORF2730
19	172	24.8	3028	21	AAK91320	Human polynucleoti
20	170.4	24.6	622	22	AAK91764	Human cDNA 5'-end
21	170.4	24.6	622	22	AAK93200	Human cDNA clone r
22	168.6	24.3	1913	22	AAH89944	Human bone marrow
23	160	23.1	1758	20	AAV84573	Human secreted pro
24	160	23.1	1758	22	ABA83356	Human secreted pro
25	132.8	19.2	1716	21	AAK77189	Human ORFX ORF3744
26	121	17.5	300	20	AAZ13026	Human gene express
27	121	17.5	710	20	AAZ15876	Human gene express
28	109.8	15.8	2837	22	AAK80493	Human immune/haema
29	104.8	15.1	483	24	ABI99871	Mouse ischaemic co
30	79.8	11.5	673	21	AAF12614	Aspergillus oryzae
31	72.4	10.4	2137	23	ABL08627	Drosophila melanog
32	67.8	9.8	3058	22	AAK83876	Human immune/haema
33	65.8	9.5	448	21	AAK77174	Human ORFX ORF2729
34	63.2	9.1	1000	21	AAA02484	Human colon cancer
35	59.2	8.5	349	22	AAI84504	Human polynucleoti
36	58.2	8.4	318	21	AAA38184	Primer used in the
37	58.2	8.4	320	21	AAA38186	Primer used in the
C	58.2	8.4	2188	20	AAZ77506	Human ovarian tumo
C	58.2	8.4	2198	20	AAZ02974	Human IL-1ra BAC c
40	58	8.4	2037	24	AAAG2787	cDNA sequence #574
41	57.6	8.3	320	21	AAA38185	Primer used in the
42	57.6	8.3	1698	23	AAAS4408	DNA encoding novel
43	57.6	8.3	2049	23	AAAS76210	DNA encoding novel
44	57.4	8.3	320	21	AAA38183	Primer used in the
45	57.2	8.3	1286	24	ABI99656	Mouse ischaemic co

ALIGNMENTS

RESULT	1
AAC90457	ID AAC90457 standard; cDNA; 1336 BP.
XX	AC AAC90457;
XX	DT 12-MAR-2001 (first entry)
XX	DE Human uncoupling protein cDNA #6.
XX	Human; uncoupling protein; immunosuppressive; antiarthritic;
KW	antirheumatic; antiproliferative; cardiant; vasotropic;
KW	cerebroprotective; neuroprotective; antibacterial; ophthalmological;
KW	gastrointestinal; nephrotropic; gynaecological; vulneryary; thrombolytic;
KW	gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
XX	infertility; ss.
XX	OS Homo sapiens.
XX	FN WO200061614-A2.
XX	PD 19-OCT-2000.
XX	PF 06-APR-2000; 2000WO-US09534.
XX	PR 09-APR-1999; 99US-0128701.
PR	08-JUL-1999; 99US-0142821.
PR	18-AUG-1999; 99US-0149448.
PR	12-NOV-1999; 99US-0164751.
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
XX	

```

PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;
XX WPI: 2000-656322/63.
DR P-PSDB: AAB50383.
XX
XX Uncoupling proteins and nucleic acid sequences encoding them, useful
PT for detecting, preventing and treating proliferative, neurological,
PT immune system, cardiovascular and gastrointestinal disorders -
XX
XX Claim 1; Page 306-307; 343pp; English.
XX
XX The present sequence is one of eighteen isolated nucleotide sequences
CC encoding uncoupling proteins. The nucleotide sequences may be used for
CC the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
CC or research purposes. The uncoupling protein encoded by the nucleotide
CC sequences may be used to treat disorders such as neural, immune,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders, wounds, infectious diseases,
CC thrombosis, arthritis, and infertility.
XX
XX Sequence 1336 BP; 333 A; 377 C; 358 G; 268 T; 0 other;

Query Match      67.7%; Score 469.4; DB 21; Length 1336;
Best Local Similarity 99.8%; Pred. No. 3.4e-91;
Matches 470; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 107 gctgtggccggggcgccgggggagcgccggggcctcagggcccccgcgtacgacaag 166
Db 6 gagtggcgggggcgccggcgccgggggagcgccggggcctcagggcccccgcgtacgacaag 65

Qy 167 atccggactccggcccgactacagcgctccggcgctggagcactgtccaccgcaca 226
Db 66 atccggactccggcccgactacagcgctccggcgctggagcactgtccaccgcaca 125

Qy 227 tgggtggcaggcgccgtggcaggatcctggagcactcgtgatcccccactgcactgcg 286
Db 126 tgggtggcaggcgccgtggcaggatcctggagcactcgtgatcccccactgcactgcg 185

Qy 287 tcaagaccggatgacagtgctacagcctgaccagctgcccgctatccgaatgtgttg 346
Db 186 tcaagaccggatgacagtgctacagcctgaccagctgcccgctatccgaatgtgttg 245

Qy 347 aggcctcttgaggattataagacgagggcctatggagcccatgaggggctgaacg 406
Db 246 aggcctcttgaggattataagacgagggcctatggagcccatgaggggctgaacg 305

Qy 407 tcacagcaacaggcgccggcctgccaccgccctttatttgcctgctacgaaaagttaa 466
Db 306 tcacagcaacaggcgccggcctgccaccgccctttatttgcctgctacgaaaagttaa 365

Qy 467 aaagacattgagtgatgaatccacctgggggccaatagccattatgccaatgtgtcgg 526
Db 366 aaagacattgagtgatgaatccacctgggggccaatagccattatgccaatgtgtcgg 425

Qy 527 ccgggtgtgtggcaacttacttctatgatgcagccatgaacccctgcgggaag 577
Db 426 ccgggtgtgtggcaacttacttctatgatgcagccatgaacccctgcgggaag 476

RESULT 2
AAC74843
ID AAC74843 standard; cdna; 1418 BP.
XX
XX AAC74843;
AC
XX
XX 08-FEB-2001 (first entry)
DT
XX
XX Human ORFX ORF398 polynucleotide sequence SEQ ID NO:795.
DE
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnerable; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW

```

```

KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621..
XX
XX 31-MAR-1999; 99US-0127607.
XX
XX 02-APR-1999; 99US-0127636.
XX
XX 05-APR-1999; 99US-0127728.
XX
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI: 2000-602362/57.
XX
XX P-PSDB; AAB40634.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 5; Page 853-854; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnerary;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antirheumatic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1418 BP; 353 A; 400 C; 368 G; 297 T; 0 other;
XX

```

```

Query Match      66.1%; Score 458; DB 21; Length 1418;
Best Local Similarity 100.0%; Pred. No. 9.3e-89;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 121 gccggcgccgggagcgccggcgccgagcccccgcgtacgacaagatccgactccgc 180
Db 1 gccggcgccgggagcgccggcgccgagcccccgcgtacgacaagatccgactccgc 60
Qy 181 ccggactacgagcgctgcggcgtggagccactgtaccacgcacatggtggcaggcgc 240

```



```
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM41505.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 4650; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 1294 BP; 267 A; 335 C; 368 G; 324 T; 0 other;

Query Match      63.5%; Score 440; DB 22; Length 1294;
Best Local Similarity 100.0%; Pred. No. 6.4e-85;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 cggggcctgagcccggtacacagatccgagtcgagccgagcagcagcctga 197
Db 1294 CGGGGCTGACAGGCGCGGTACACAGATCCGGACTCCGGCCCGGACTACGAGGGCGCT 1235
QY 198 gccgctggagccactgcaccagcacatggtgagcagcgcctgagcagggatcctgga 257
Db 1234 CGCGGCTGGAGCCACTGTACACGACACATGTTGGCAGGCGCCGTGGCAGGGATCCTGGA 1175
QY 258 gcactgctgatgtaccgccatcgactcgctcaagaccgagtcagagctctacagcctga 317
Db 1174 GCACCTGCTGATGTACCCATCGACTCGTCAAGACCCGGATGTCAGAGCTCTACAGCCTGA 1115
QY 318 ccaagctgcgcgctatgcgaatgtgttggaggccctctgagagattataagaacgaggg 377
Db 1114 CCCAGCTGCCCGCTATCCGAATGTGTGGAGGCCCTCTGGAGGATTTAAGACAGCGAGG 1055
QY 378 cctatgagcccatgaggggctgaacgtcacagcaacagcagcagcgcctgccacgc 437
Db 1054 CCTATGAGGCCCATGAGGGGCTGAACGTCTACAGCAACAGGCGCAGGGCTGCCCAACG 995
QY 438 cctttatttgcctgctacgaaagttaaaagacacattgagtgatgtaattccacctgg 497
Db 994 CCTTTATTTTGGCTGCTACGAAAGTTAAAAAGACATTGAGTGATGTAATCCACCTGG 935
QY 498 gggcaatagcatattccaatggtgcggcggggtgtgtggcaacattatctatgatgc 557
Db 934 GGGCAATAGCATATTCCCAATGTGGCGGGGTGTGTGCAACATTACTTATGATGC 875
QY 558 agccatgaacccctgcgaag 577
Db 874 AGCCATGAACCCCTGCGGAAG 855
```

```
RESULT      5
AAS03906
ID AAS03906 standard; cDNA; 2502 BP.
XX
XX AAS03906;
XX
XX 29-AUG-2001 (first entry)
XX
XX Human secreted protein gene #25.
XX
XX Human secreted protein; autoimmune disorder; hyperproliferative disorder;
XX cardiovascular disorder; cerebrovascular disorder; angiogenesis;
XX nervous system disorder; bacterial infection; viral infection; ss;
XX fungal infection; ocular disorder; wound healing; tissue regeneration;
XX epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.
XX
XX Homo sapiens.
XX
XX WO2001233598-A1.
XX
XX 05-APR-2001.
XX
XX 26-SEP-2000; 2000WO-US26324.
XX
XX 27-SEP-1999; 99US-0155807.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Komatsoulis G, Ruben SM, Rosen CA;
XX
XX WPI: 2001-281684/29.
XX P-PSDB; AAU01950, AAU01989.
XX
XX Forty one nucleic acid molecules encoding human secreted proteins, useful
XX in the prevention, treatment and diagnosis of cancer, immune disorders,
XX cardiovascular disorders and neurological diseases -
XX
XX Disclosure; Page 460-461; 518pp; English.
XX
XX Sequences AAS03873-AAS03922 represent isolated nucleic acid molecules and
XX PCR primers of the invention. acid of the invention. Secreted proteins
XX and their related nucleic acids can be used in the diagnosis of or
XX susceptibility to a pathological condition by determining the presence or
XX absence of a mutation in a nucleic acid or the presence or amount of
XX expression of a secreted protein. The sequences are used to prevent,
XX treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
XX goats, horses, cats, dogs, chickens or sheep. The antibodies to the
XX polypeptides can also be used in alleviating symptoms associated with
XX disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
XX linked immunosorbent assays (ELISA). The disorders include autoimmune
XX diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
XX neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
XX arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
XX nervous system disorders e.g. Alzheimer's disease, infections caused by
XX bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
XX The peptides can also be used to aid wound healing and epithelial cell
XX proliferation, to help prevent skin ageing due to sunburn, to maintain
XX organs before transplantation, to regenerate tissues, in chemotaxis and
XX as a food additive or preservative to alter storage capabilities.
XX
XX Sequence 2502 BP; 587 A; 655 C; 669 G; 589 T; 2 other;
```

```
Query Match      57.9%; Score 401; DB 22; Length 2502;
Best Local Similarity 99.5%; Pred. No. 1.5e-76;
Matches 412; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 109 gtgggcccggggccgagcgcgagggagccggccgctgcagggcccggtacgacaagat 168
Db 7 gtgggcccggggccgagcgcgagggagccggggccgctgcaggg-cctgggtacgacaagat 65
QY 169 ccggactccggcccgactacgagcgcgctgcggctggagccactgtcaccacgcacatg 228
```


Db 66 cggagactcggccgactacagagcgctgcggcctggagccaactgtccaccgacacatg 125
QY 229 gtgcagggccgctggcaggatcctggagcactgcgtgatgtaccccatcgactcgctc 288
Db 126 gtgcagggccgctggcaggatcctggagcactgcgtgatgtaccccatcgactcgctc 185
QY 289 aagaccggatgcagatgtacagctgacccagctgcccgcctatgcgaatgtgtggag 348
Db 186 aagaccggatgcagatgtacagctgacccagctgcccgcctatgcgaatgtgtggag 245
QY 349 gccctctgagagattataaagcaggaggccctatgaggcccatgagggggctgaacgtc 408
Db 246 gccctctgagagattataaagcaggaggccctatgaggcccatgagggggctgaacgtc 305
QY 409 aagcaacagggcgagggcctgcgccacgccccttatttgcctgctacgaaagttaaaa 468
Db 306 acagcaacagggcgagggcctgcgccacgccccttatttgcctgctacgaaagttaaaa 365
QY 469 aagacattgagtgtatccaccctggggcaatagccatattgccaatggt 522
Db 366 aagacattgagtgtatccaccctggggcaatagccatattgccaatggt 419

RESULT 6
AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX
AC AAI58875;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1078.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.

XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB; AAM39719.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1: SEQ ID NO 1078; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;

Query Match 57.5%; Score 398.6; DB 22; Length 1316;
Best Local Similarity 95.5%; Pred. No. 4.4e-76;
Matches 421; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 138 cggggcctcaggccccggtacgacaatccggactccggcccgactacagcgct 197
Db 15 cgaggcctcgagggccccggtacgacatccgactccggcccgactacagcgct 74
QY 198 gccggctgagccactgtcaccacgcacatg-gtggcaggcgccgtggcaggactcctgg 256
Db 75 gtccgctgagccactgtcaccacgcacatgtgggtatgctgcgcgagccatggtcgtg 134
QY 257 agcactgcgtgatgtaccccatcgactgcgtcaagaccggatgcagatctacagcctg 316
Db 135 agcactgcgtgatgtaccccatcgactgcgtgatgaccggatgcagatctacagcctg 194
QY 317 acccagctcccgctatcgcaatgtgtggaggccctctggaggattataagaaacggagg 376
Db 195 acccagctggcgctatcgcaatgtgtggaggctctctggaggattataagaaacggagg 254
QY 377 gcctatggagcccatgagggggctgaacgtcacacacagcagcgccgtgcccacg 436
Db 255 gcctatggagggcccatgagggggctgaacgtcacacacagcagcgccgtgcccacg 314
QY 437 cccttattttgcctgtctacgaaaagttaaaaaagacattgagtgttaatccaccctg 496
Db 315 cccttattttgcctgtctacgaaaagttaaaaaagacattgagtgttaatccaccctg 374
QY 497 ggggcaatagccaattgtccaatgtgtgcggccgggtgtgtggcaacattacttcattg 556
Db 375 ggggcaatagccaattgtccaatgtgtgcggccgggtgtgtggcaacattacttcattg 434
QY 557 cagccatgaacctgcggaag 577
Db 435 cagccatgaacctgcggaag 455

RESULT 7
AAF66156
ID AAF66156 standard; cDNA; 401 BP.
XX
AC AAF66156;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1912.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX

PR	14-SEP-2000;	2000US-0233066;
PR	14-SEP-2000;	2000US-0233066;
PR	21-SEP-2000;	2000US-0234232;
PR	21-SEP-2000;	2000US-0234232;
PR	21-SEP-2000;	2000US-0234274;
PR	25-SEP-2000;	2000US-0234597;
PR	25-SEP-2000;	2000US-0234597;
PR	25-SEP-2000;	2000US-0234998;
PR	26-SEP-2000;	2000US-0234998;
PR	27-SEP-2000;	2000US-0235484;
PR	27-SEP-2000;	2000US-0235484;
PR	27-SEP-2000;	2000US-0235836;
PR	27-SEP-2000;	2000US-0235836;
PR	29-SEP-2000;	2000US-0236327;
PR	29-SEP-2000;	2000US-0236367;
PR	29-SEP-2000;	2000US-0236368;
PR	29-SEP-2000;	2000US-0236368;
PR	29-SEP-2000;	2000US-0236370;
PR	29-SEP-2000;	2000US-0236370;
PR	02-OCT-2000;	2000US-0236802;
PR	02-OCT-2000;	2000US-0236802;
PR	02-OCT-2000;	2000US-0237037;
PR	02-OCT-2000;	2000US-0237038;
PR	02-OCT-2000;	2000US-0237039;
PR	02-OCT-2000;	2000US-0237043;
PR	13-OCT-2000;	2000US-0239935;
PR	13-OCT-2000;	2000US-0239935;
PR	13-OCT-2000;	2000US-0239937;
PR	20-OCT-2000;	2000US-0240960;
PR	20-OCT-2000;	2000US-0240960;
PR	20-OCT-2000;	2000US-0241221;
PR	20-OCT-2000;	2000US-0241785;
PR	20-OCT-2000;	2000US-0241785;
PR	20-OCT-2000;	2000US-0241786;
PR	20-OCT-2000;	2000US-0241787;
PR	20-OCT-2000;	2000US-0241808;
PR	20-OCT-2000;	2000US-0241809;
PR	20-OCT-2000;	2000US-0241826;
PR	01-NOV-2000;	2000US-0244517;
PR	08-NOV-2000;	2000US-0246474;
PR	08-NOV-2000;	2000US-0246474;
PR	08-NOV-2000;	2000US-0246475;
PR	08-NOV-2000;	2000US-0246476;
PR	08-NOV-2000;	2000US-0246477;
PR	08-NOV-2000;	2000US-0246478;
PR	08-NOV-2000;	2000US-0246523;
PR	08-NOV-2000;	2000US-0246523;
PR	08-NOV-2000;	2000US-0246524;
PR	08-NOV-2000;	2000US-0246525;
PR	08-NOV-2000;	2000US-0246526;
PR	08-NOV-2000;	2000US-0249210;
PR	17-NOV-2000;	2000US-0249211;
PR	08-NOV-2000;	2000US-0246613;
PR	17-NOV-2000;	2000US-0249207;
PR	17-NOV-2000;	2000US-0249208;
PR	17-NOV-2000;	2000US-0249209;
PR	17-NOV-2000;	2000US-0249210;
PR	17-NOV-2000;	2000US-0249211;
PR	17-NOV-2000;	2000US-0249212;
PR	17-NOV-2000;	2000US-0249213;
PR	17-NOV-2000;	2000US-0249213;
PR	17-NOV-2000;	2000US-0249216;
PR	17-NOV-2000;	2000US-0249216;
PR	17-NOV-2000;	2000US-0249215;
PR	17-NOV-2000;	2000US-0249216;
PR	17-NOV-2000;	2000US-0249265;
PR	17-NOV-2000;	2000US-0249267;
PR	17-NOV-2000;	2000US-0249297;
PR	17-NOV-2000;	2000US-0249299;
PR	17-NOV-2000;	2000US-0249300;
PR	01-DEC-2000;	2000US-0250160;
PR	01-DEC-2000;	2000US-0250391;
PR	05-DEC-2000;	2000US-0251030;
PR	05-DEC-2000;	2000US-0251988;
PR	05-DEC-2000;	2000US-0251988;
PR	06-DEC-2000;	2000US-0256719;
PR	06-DEC-2000;	2000US-0256719;
PR	08-DEC-2000;	2000US-0251856;
PR	08-DEC-2000;	2000US-0251856;

	PR	08-DEC-2000; 2000US-0251869.
	PR	08-DEC-2000; 2000US-0251989.
	PR	08-DEC-2000; 2000US-0251990.
	PR	11-DEC-2000; 2000US-0254097.
	PR	05-JAN-2001; 2001US-0259678.
	XX	(HUMA-) HUMAN GENOME SCI INC.
	PA	Rosen CA, Barash SC, Ruben SM;
	PI	WPI; 2001-483426/52..
	DR	
	XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
	PT	useful for preventing, diagnosing and/or treating cancers and
	PT	metastasis -
	XX	
	PS	Disclosure; SEQ ID NO 24591; 307lpp + Sequence Listing; English.
	XX	
	CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
	CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
	CC	activity, and can be used in gene therapy and vaccine production. (I)
	CC	proteins and polynucleotides may be used in the prevention, diagnosis and
	CC	treatment of diseases associated with inappropriate (I) expression. For
	CC	example, they may be used to treat disorders associated with decreased
	CC	expression by rectifying mutations or deletions in a patient's genome
	CC	that affect the activity of (I) by expressing inactive proteins or to
	CC	supplement the patients own production of (I). Additionally, (I)
	CC	polynucleotides may be used to produce the secreted (I), by inserting
	CC	the nucleic acids into a host cell and culturing the cell to express the
	CC	protein. (I) proteins and polynucleotides may be used to prevent,
	CC	diagnose and treat immune/haematopoietic-related diseases, especially
	CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703
	CC	to AAK87694 represent human immune/haematopoietic antigen genomic
	CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169
	CC	represent sequences used in the exemplification of the present invention.
	XX	
	SQ	Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other:
		Query Match 33.6%; Score 233; DB 22; Length 27960;
		Best Local Similarity 94.2%; Pred. No. 1.4e-40;
		Matches 24; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
	QY	266 tgatgtaccccatgcagtctgctcaagaccggatgccagtgctcagcgacctgaggactgtg
	DB	18263 tgcgttcttcgttgttaattcacagcccggatgcagttctaagctgaagccccgatg
	QY	326 ccgcgtatcgaaatgtgttgaggccctctggaggattataagaacggagggcctatgga
	DB	18323 cccgctatcgaaatgtgttgaggccctctggaggattataagaacggagggcctatgga
	QY	386 ggcccatgaggggctgaacgtccaagacaacagcgcaggccctgccacgcccttatatt
	DB	18383 ggcccatgaggggctgaacgtccaagacaacagcgcaggccctgccacgcccttatatt
	QY	446 ttgcctctcagaaaagttaaanaaacattgagtgatgaatccaccctgggggcaata
	DB	18443 ttgcctctcagaaaagttaaanaaacattgagtgatgaatccaccctgggggcaata
	QY	506 gccatatggcaaattggt 522
	DB	18503 gccatatggcaaattggt 18519
	RESULT 9	
	AAK73320	ID AAK73320 standard; DNA; 27960 BP.
	XX	AAK73320;
	AC	
	XX	
	DT	06-NOV-2001 (first entry)
	XX	
	XX	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28132.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX WO200157182-A2.
PN
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233053.
PR 14-SEP-2000; 2000US-0233054.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-02334223.
PR
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.

PR	02-MAR-2000;	2000US-01863550;
PR	16-MAR-2000;	2000US-01989874;
PR	17-MAR-2000;	2000US-01900076;
PR	18-APR-2000;	2000US-01981123;
PR	19-MAY-2000;	2000US-02055515;
PR	07-JUN-2000;	2000US-02094677;
PR	28-JUN-2000;	2000US-02144866;
PR	30-JUN-2000;	2000US-02151335;
PR	07-JUL-2000;	2000US-02166647;
PR	07-JUL-2000;	2000US-02165880;
PR	11-JUL-2000;	2000US-02174877;
PR	11-JUL-2000;	2000US-02174966;
PR	14-JUL-2000;	2000US-02182890;
PR	26-JUL-2000;	2000US-02209633;
PR	26-JUL-2000;	2000US-02209664;
PR	14-AUG-2000;	2000US-02245418;
PR	14-AUG-2000;	2000US-02245419;
PR	14-AUG-2000;	2000US-02252213;
PR	14-AUG-2000;	2000US-02252211;
PR	14-AUG-2000;	2000US-02252266;
PR	14-AUG-2000;	2000US-02252567;
PR	14-AUG-2000;	2000US-02252568;
PR	14-AUG-2000;	2000US-02253477;
PR	14-AUG-2000;	2000US-02254240;
PR	14-AUG-2000;	2000US-02257577;
PR	14-AUG-2000;	2000US-02257578;
PR	14-AUG-2000;	2000US-02277009;
PR	23-AUG-2000;	2000US-02277182;
PR	30-AUG-2000;	2000US-02289924;
PR	01-SEP-2000;	2000US-02292877;
PR	01-SEP-2000;	2000US-02293433;
PR	01-SEP-2000;	2000US-02293444;
PR	01-SEP-2000;	2000US-02293445;
PR	05-SEP-2000;	2000US-02295909;
PR	05-SEP-2000;	2000US-02295913;
PR	06-SEP-2000;	2000US-02304337;
PR	06-SEP-2000;	2000US-02304338;
PR	08-SEP-2000;	2000US-02312422;
PR	08-SEP-2000;	2000US-02312423;
PR	08-SEP-2000;	2000US-02312444;
PR	08-SEP-2000;	2000US-02314133;
PR	08-SEP-2000;	2000US-02314134;
PR	08-SEP-2000;	2000US-02320880;
PR	08-SEP-2000;	2000US-02320881;
PR	12-SEP-2000;	2000US-02331968;
PR	14-SEP-2000;	2000US-02330663;
PR	14-SEP-2000;	2000US-02330664;
PR	14-SEP-2000;	2000US-02330665;
PR	21-SEP-2000;	2000US-02342223;
PR	21-SEP-2000;	2000US-02342274;
PR	25-SEP-2000;	2000US-02349999;
PR	25-SEP-2000;	2000US-02349998;
PR	26-SEP-2000;	2000US-02354884;
PR	27-SEP-2000;	2000US-02356834;
PR	27-SEP-2000;	2000US-02358636;
PR	29-SEP-2000;	2000US-02363232;
PR	29-SEP-2000;	2000US-02363667;
PR	29-SEP-2000;	2000US-02363668;
PR	29-SEP-2000;	2000US-02363669;
PR	29-SEP-2000;	2000US-02363700;
PR	02-OCT-2000;	2000US-02368602;
PR	02-OCT-2000;	2000US-02370037;
PR	02-OCT-2000;	2000US-02370038;
PR	02-OCT-2000;	2000US-02370039;

SQ Sequence 2562 BP; 623 A; 660 C; 627 G; 652 T; 0 other;

Query Match 27.0%; Score 187.4; DB 22; Length 2562;
Best Local Similarity 68.2%; Pred. No. 5.6e-31;

Best local similarity	90.26	File: NO: 3.06	31
Matches	260	Conservative	0
Matches	260	Mismatches	121
		Indels	0
		Gaps	0

Qy 197 tgcggctggagccactgtcaccacgcacatggtggtgcaggcgcgcgtgcagggatcctgg 256

Db 1920 tgccgactagcgcctccggtccaccacatgacagcaggagcgatgccggggatccctgg 1979

Qy 257 agcactgcgtgatgtaccocatgcactgcgtcaagacccggatgcagaggtctacagcctg 316

Db 1980 agcactcgggtcatgtaccacggtggactcgggtgaagacacgaaatgcagagttttgagtcacg 20399

Qy 317 accagctgcgcgtatcgcaatgtgttgaggccctctggagattataagaacgagg 376

Db 2040 atccaaagcccgtagacaagtatctacggagccctcaagaaataatcatgaggacgaag 2099

Qy 377 gcctatggagcccatgagggggtgaacgtcacagaacagcgagcgctgccacg 436

Db 2100 gctctggaggcccttgaggcgtcaacgcatgatcatgggtgcaggcgccgcctg 2159

Qy 437 ccctttatttgcctgctacgaaaagttaaaaaagacattgagtgatgtaatccaccctg 496
||| |||||||||||||||| ||||| | ||||| ||||| ||||| ||||| ||||| |||||
Pb 2150 cccctttatttgcctgctacgaaaagttaaaaaagacattgagtgatgtaatccaccctg 2210
||| |||||||||||||||| ||||| | ||||| ||||| ||||| ||||| ||||| |||||

D_B 2160 ccattgatttttgctgctatgaacaacatgaaaaggacitaaatgacgttttccaccacc 2219
Q_U 407 gggggaaataagcgcataatgcccaatatcccccacccatgacgcatcgatgac 555

pQ 99gggcaatagccatatctgccaatgtgcggcggggtgtgtggcaacttacttcacgatg 356
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
pB 2220 aaaggaaaccccgaacctaccacaacggatatacctgggaattatggcacccccctgcaccaactn 378

0v 557 cagccatgaacctcgcgaag 577
 zzz0 aagggaacagccacctagccacgggatagctgggagatggccacccctgctccacgatg 22/

QY 337 cagccatgaacccccgcggaag 377
| | | | |
pb 2280 caataatgaatccagcagaag 2300

RESULT 13

RESOLUT 13
AAK58672

ID AAK58672 standard; cDNA; 339 BP.

XX
XX

AC AAK58672;

XX
DE
100-110630
FBI
C + WDT 06-NOV-2001 (first entry)
XX

Human immune/haematopoietic anti-

XX ;
XX

KW Human; immune; haematopoietic; i

KW cytostatic; gene therapy; vaccin

XX

OS Homo sapiens.

XX
DNI
W0300157193-A3

PN
XX
W0200157182-A2.

XX
PD
09-AUG-2001

XX
XX
09 DEC 2001.

PF 17-JAN-2001; 2001WO-US01354.

XX
XX

PR 31-JAN-2000; 2000US-0179065.

PR 04 -FEB-2000; 2000US-0180628.

PR 24 -FEB-2000; 2000US-0184664.

P-PSDB; AAK93934.
 830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 4109; 1380pp + sequence listing; English.
 XX
 The invention relates to primers for synthesising full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesising the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 Sequence 1890 BP; 441 A; 534 C; 496 G; 419 T; 0 other;
 SQ

Query Match 25.1%; Score 174.2; DB 22; Length 1890;
 Best Local Similarity 62.4%; Pred. No. 3.6e-28;
 Matches 289; Conservative 0; Mismatches 173; Indels 1; Gaps 1

QY 61 cggagcccgaggagtcggcgtctgacggggggtgctgacggcggtggtgcccggggg 120
 Db 26 ctgctgcccgcgagctggtggtgagctgcagcggggcgctggcagccagcggt 85
 QY 121 gccgcgcggggaggccggcggtcgagggcccccggtacgacaaagtccggactccg 180
 Db 86 gccgcggaggtggtgggacagccgagatggcggcgcggaagacgccaccgggtc 145
 QY 181 -ccggactacgggcgtgcgcggctggagccactgtccaccatgcacatgtggcagcgc 239
 Db 146 agaggactacgagaacctgcgactagcgtccgtggtccaccacatgacagcaggagc 205
 QY 240 cgtgcgaggtacctggagcactgcgtatgtaccctcagctgcacagaccggat 299
 Db 206 gatggcggggatccctggagcactcgtgtatcccggtggagctcgtggaacacgat 265
 QY 300 gcagagtctacagcctgaccacgtgcccgctatcgcaatgtgtgaggccctctggag 359
 Db 266 gcagatttgatccagatcccaagccacgtacacagtatctacggagccctcaga 325
 QY 360 gattataagaacggaggccctatggaggcccatgaggggcgctgaactcagcacaacagg 419
 Db 326 aatcatcggaacgaaggctcttgaggcctctgcaggcgtcaactgcatgcatggg 385
 QY 420 cgcaggcctgccacgcccttatttgcctgtcagaaagttaaaagacattgag 479
 Db 386 tgcaggcgcggcccatgcattgtatttgcctgtatgaaacatgaaaggactttaa 445
 QY 480 tgatgtaatccaccctggggcaatagccatttgccaatggt 522
 Db 446 tgacgtttccaccaccaggaaacagccactagccaacggt 488

RESULT 15
 AAK92357
 ID AAK92357 standard; cDNA; 842 BP.
 XX AC AAK92357;
 XX
 XX 06-NOV-2001 (first entry)
 XX
 XX Human cDNA 5'-end sequence, SEQ ID NO: 817.
 DE
 XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 KW
 XX Homo sapiens.
 OS

Search completed: September 28, 2002, 08:05:59
Job time: 11554 sec

Homo sapiens.

Result No.	Score	Query Match	Length	DB	ID	Description	
1	62.4	9.0	319	4	US-09-165-264-8	Sequence 8, Appl	
2	58.2	8.4	318	4	US-09-165-264-12	Sequence 12, Appl	
3	58.2	8.4	320	4	US-09-165-264-14	Sequence 14, Appl	
C 4	58.2	8.4	152331	3	US-09-128-155-16	Sequence 16, Appl	
	57.8	8.3	320	4	US-09-165-264-7	Sequence 7, Appl	
6	57.6	8.3	320	4	US-09-165-264-13	Sequence 13, Appl	
7	57.4	8.3	320	4	US-09-165-264-11	Sequence 11, Appl	
C 8	50.6	7.3	833	2	US-08-403-852D-3	Sequence 3, Appl	
C 9	50.6	7.3	833	3	US-08-510-646B-3	Sequence 3, Appl	
C 10	50.6	7.3	833	4	US-09-231-818-3	Sequence 3, Appl	
C 11	50.6	7.3	5392	2	US-08-403-852D-1	Sequence 1, Appl	
C 12	50.6	7.3	5392	3	US-08-510-646B-1	Sequence 1, Appl	
C 13	50.6	7.3	5392	4	US-09-231-818-1	Sequence 1, Appl	
14	49.4	7.1	43280	2	US-08-804-227C-1	Sequence 1, Appl	
15	47	6.8	1227	2	US-08-903-800A-1	Sequence 1, Appl	
C 16	46.8	6.8	1931	2	US-09-130-114-2	Sequence 2, Appl	
C 17	46.8	6.8	4257	2	US-08-690-473-1	Sequence 1, Appl	
C 18	46.8	6.8	4257	4	US-09-259-821A-1	Sequence 1, Appl	
C 19	46.8	6.8	4257	4	US-08-843-659-1	Sequence 1, Appl	
	46.8	6.8	12001	1	US-08-458-568A-11	Sequence 11, Appl	
20	46.6	6.7	4403765	4	US-09-103-840A-2	Sequence 2, Appl	
21	46.6	6.7	4411529	4	US-09-103-840A-1	Sequence 1, Appl	
22	46.6	6.7	4411529	4	US-09-103-840A-2	Sequence 2, Appl	
C 23	46.2	6.7	2823	1	US-08-398-008A-1	Sequence 1, Appl	
C 24	46.2	6.7	2823	2	US-08-393-333-1	Sequence 1, Appl	
25	46.2	6.7	3415	1	US-08-894-077C-1	Sequence 1, Appl	
26	45.8	6.6	460	2	US-08-903-800A-4	Sequence 4, Appl	
C 27	45.6	6.6	1300	4	US-08-483-533-39	Sequence 39, Appl	
	45.6	6.6	1300	4	US-08-483-533-39	Sequence 39, Appl	

```
; SEQ ID NO 12
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-12

Query Match      8.4%; Score 58.2; DB 4; Length 318;
Best Local Similarity 61.6%; Pred. No. 5.1e-05;
Matches 93; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 3 ggaattgagggcggtgctgctggtgagcggggcgccgagcagggcccgagcg 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 gggggggggggggggggggggggggggggggggggggggggggggggggg 226

Qy 63 gagccccggggagtcgcgtgctgagcgggtggtgctcagcggggcgggggc 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 gggggggggggggggggggggggggggggggggggggggggggggggggg 286

Qy 123 cggcgcgggggagggcgccgctcagggccc 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 ggggggggggggggtggtggtgtaagccc 317

RESULT 3
US-09-165-264-14
; Sequence 14, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamoorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-14

Query Match      8.4%; Score 58.2; DB 4; Length 320;
Best Local Similarity 50.5%; Pred. No. 5.1e-05;
Matches 141; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 3 ggaattgagggcggtgctgctggtgagcggggcgccgagcagggcccgagcg 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 42 gggggggggggggggggggggggggggggggggggggggggggggggggg 101

Qy 63 gagccccggggagtcgcgtgctgagcgggtgctcagcggggcggtgagcg 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 gggggggggggggggggggggggggggggggggggggggggggggggggg 161

Qy 123 cggcgcgggggagggcgccgctcagcccccggtacgacagatccgactccgccc 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 gggggggggggggggggggggggggggggggggggggggggggggggggg 221

Qy 183 ggaactagagcgctgcggctgagcactgtcacacgcacatggtgcagggcgct 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 gggggggggggggggggggggggggggggggggggggggggggggggggg 281

Qy 243 ggcagggatcctcgagcactgcgtgattaccatccatcga 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 gggggggggggggggggggggggggggggggggggggggggggggggggg 320

RESULT 4
```

```
US-09-128-155-16/c
; Sequence 16, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

Query Match      8.4%; Score 58.2; DB 3; Length 152331;
Best Local Similarity 55.1%; Pred. No. 0.00018;
Matches 114; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy 8 tggagggggcggtgctgctggtgagcggggcgccgagcagggcccgagcgcc 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22293 TCCCGGTCGGGGCGGGGGGGGCTGGGGGGGGTGGGGGGGGCGCGGTCGGGG 22234

Qy 68 ccgggggagtcggcgctgctgagcgggtggtgctgagcggggcggtgggcg 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22233 CTGGGGGGGGGGGTGGGGGGCGCGCGGGGGGGGGGGGGGGGGGGGGGGG 22174

Qy 128 gcgggagcgccggcgctcagcccccggtacgacagatccgactccggcgact 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22173 GCGGGCGGGGGGGGGGTGCGGGGTGCGGGGGGGGGGGGGGGGGGGGGGGG 22114

Qy 188 acgagggcgctccgggtggagccactg 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22113 GCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 22087

RESULT 5
US-09-165-264-7
; Sequence 7, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamoorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-7

Query Match      8.3%; Score 57.8; DB 4; Length 320;
Best Local Similarity 61.7%; Pred. No. 6.3e-05;
Matches 92; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 06:22:39 ; Search time 4881.77 Seconds
(without alignments)
1915.983 Million cell updates/sec

Title: US-09-870-113-5
Perfect score: 693
Sequence: 1 atgagtgtagggcgggg.....ttacctgtttaccactag 693

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	532.6	76.9	888	9	AL530804
2	450.2	65.0	817	10	BI103329
3	438.4	63.3	1131	10	BG295496
4	419.4	60.5	573	10	BI541703
5	409.8	59.1	697	10	BI854643
6	398.2	57.5	967	10	BI739710
7	386.6	55.8	645	10	BG964218
8	367.2	53.0	397	9	BE012485
9	354.6	51.2	574	10	BI185580
10	346.4	50.0	681	9	AV704087
11	344.6	49.7	598	10	BE913718
12	323	46.6	560	10	BI794616
13	320.2	46.2	480	10	BI045863
14	302.2	43.6	336	9	AW326482
15	298.4	43.1	546	10	BM488747
16	292	42.1	546	9	BE014800
17	290.8	42.0	900	10	BG325430

18	288.6	41.6	747	10	BI827322
19	282.8	40.8	501	9	AW211366
20	274.4	39.6	916	10	BI757180
21	255.8	36.9	622	10	BJ039504
22	244.6	35.3	722	10	BI876051
23	243.6	35.2	417	9	AA104365
24	240.4	34.7	855	10	BI412175
25	240.2	34.7	261	10	BI338246
26	231.4	33.4	701	10	BI831943
27	222.6	32.1	421	10	BI448355
28	207.4	29.9	678	9	BG654057
29	203.2	29.3	561	9	AI942584
30	196.6	28.4	684	9	AI133696
31	196.6	28.4	750	10	BI831750
32	194.6	28.1	729	10	BI460337
33	194	28.0	567	10	BJ035232
34	193.6	27.9	512	9	AW210205
35	191.6	27.6	453	10	BF916224
36	190.2	27.4	612	10	BJ012200
37	187.4	27.0	1295	11	AK015770
38	186.8	27.0	803	10	BG777404
39	184.6	26.6	217	9	AW346915
40	182.8	26.4	331	10	BI449976
41	181.6	26.2	1209	11	AK006155
42	177.8	25.7	666	9	BB644566
43	177.2	25.6	513	10	BI466991
44	174.2	25.1	866	10	BF797743
45	172	24.8	592	10	BM313240

ALIGNMENTS

RESULT 1

AL530804

LOCUS

DEFINITION

AL530804

ACCESSION

AL530804

VERSION

AL530804.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 888)

AUTHORS

Li.W.B., Gruber.C., Jessee.J. and Polayes.D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..888

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DD008YH17"

/clone_lib="LTI_NFL001_NBC4"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with NotI and

cloned into the NotI and EcoRV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 168 a 244 c 322 g 139 t 15 others

AL530804 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD008YH17 5
prime, mRNA sequence.

AL530804 888 bp mRNA linear EST 13-FEB-2001

AL530804.1 GI:12794297

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li.W.B., Gruber.C., Jessee.J. and Polayes.D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..888

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DD008YH17"

/clone_lib="LTI_NFL001_NBC4"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with NotI and

cloned into the NotI and EcoRV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 168 a 244 c 322 g 139 t 15 others

ORIGIN

Query Match 76.9%; Score 532.6; DB 9; Length 888;

Best Local Similarity 98.1%; Pred. No. 1.le-93;

Matches 567; Conservative 3; Mismatches 4; Indels 4; Gaps 3;

Qy 1 atggagctggagggcgccggtgctgagcgtgtgagcggggcgccgagcggcgccggg 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 122 ATGGAGTAGAGGGCGGGGCTGCGCGTGTGGCGGGGGCGCGCGAGGCCCGGG 181
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 61 cggagcccccggggagtcgctgctgacggtgtgctgcagcggggtgctggcgggg 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 182 CGGAG--CGGGGAGTCGCGCTGCTGACGCGGTGCTGACGCGGGCGTGGCGCGGGG 239
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 121 gccgagcgggggagcggggcctgagcccccggcctgacagaagatccggactccggc 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 240 GCCGCGCGGGGAGCGCGGGGCTGCGAGG--CCCCGGTACGCAAGATCCGGACTCCGGC 298
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 181 ccggactacgagcgtgctccgg--ctggagccactgtcaccacacatggtgagcggc 239
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 299 CCGGACTACGAGGCGCTGCCGNCCTGGAGCCACTGTCTACCACGCACATGTTGGCAGGCGC 358
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 240 cgtgagggatcctgagcactgctgctgctgctgctgctgctgctgctgctgctgct 299
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 359 CGTGCGAGGGATCTCTGGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 300 gcagagctacagcctgaccagcgtgcccgtctatgcagcaatgtgtggagggccttgag 359
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 419 GCAGAGTCTACAGCCTGACCCAGCTGCCCGCTATCGCAATGTTGGAGGCGCTTGAG 478
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 360 gattataagacggagggcctatgagggccatgaggggctgaacgtcacagcaacagg 419
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 479 GATTATAGACGGAGGCGCTTATGGAGGCGCCATGAGGGGGCTGACGTCACAGCAGCAGG 538
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 420 cgcagggcctgccacgcccctttatttgcctgctacgaaagttaaaaaagacattgag 479
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 539 CCGAGGGCTGCCACGCGCTTATTTTTCCTGCBACGAAAGTTAAAAAGACATTGAG 598
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 480 tgaatgaatcaccctgggggcaatagcattatgcaaatggtgagcggggtgtgagc 539
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 599 TGAATTTATCCACCTGGGGGCAATAGCAATATTGCCAATGTTGGCGGGGTGTGTGGC 658
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 540 aacattactcatgctgagcgcattgaacccctgcggaag 577
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 659 AACATTACTCATGATGAGCCATGACCCCTGCGGAG 696
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 2

BII03329
LOCUS BII03329 817 bp mRNA linear EST 26-JUN-2001
DEFINITION 60288838F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5043992
5', mRNA sequence.

ACCESSION BII03329

VERSION BII03329.1 GI:14554222

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 817)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM1120 row: 1 column: 09

High quality sequence stop: 777.

FEATURES
source

1. 817
Location/Qualifiers
/Organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/Clone="IMAGE:5043992"
/clone_lib="NCI_CGAP_Kid14"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"

BASE COUNT 147 a 222 c 316 g 132 t
ORIGIN

Query Match 65.0%; Score 450.2; DB 10; Length 817;

Best Local Similarity 90.9%; Pred. No. 1e-77;

Matches 501; Conservative 0; Mismatches 48; Indels 2; Gaps 2;

Qy 29 gtgtggcgggggcgccgagggcgccggcgagcccgggagtcgctgctgtg 88
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 153 GTGCTGGAGTGACCATGCTGCTGGCGCGCGGAGCCCGGGGAGTCGGCGCTGCTGG 212
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 89 acgggtgctcagcgggctggtggcgggcgggcgggcgggcgggcgggcgggcgctgca 148
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 213 ACGGCTGGCTCAGCGGGGCTGGCGGGGGCGGGCGGGGAGGCGGGGCGCTATC 272
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 149 ggcggcggtgcagcaagatccgactccgcccggactacgagcgtgcgcgctgag 208
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 273 AGCCCCCTGTACGGCTGGATCCGGAGTCCGGCCCGAATACGAGCGTGCCTGGAG 332
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 209 ccactgtcaccacgcacatggtgagcggcgccgtgagggatcctgagcactgctga 268
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 333 CCACTGTCAACCGCACATGTTGGCGCGCGCTGGCAGGATCTTGGAGCATTCGCTGA 392
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 269 tgtacccatcagcgtcgaagaccgagatcgagagctacagcctgacagcagtcgcc 328
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 393 TGTACCCGATCGACTGCTCAAGACCCGGATGCAGAGCCTACAGCCTACCCAGCGGCC 452
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 329 gctatcgcaatgtgtgagccctctgaggtattataagaacgagagccctatgagcc 388
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 453 GCTATCGGAACGTGTGGAGGCTCTCTGAGAAATCATCAGGACAGAGGCGCTGTGAGGC 512
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 389 ccattgaggggctgaagtcacagcagcagcagcagcagcagcagcagcagcagcagc 448
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 513 CCATCGGGGGCTGAACGTCACAGCACAGCGGGGGCTGCCACGCGCTCTATTATTG 572
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 449 cctgctacgaaagttaaaagacattgagtgatgtaataccaccctgggggcaatagcc 508
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 573 CTTGCTAGAAAAGTTAAAAAGACATTGAGTGAGTAATCCACCCAGGGGGCAATAGCC 632
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 509 atattgccaatgg-tgcggcgccgggtgtgtggcacaattac-ttcatgatgcagccatgaa 566
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 633 ATATTGCAATGGTTGCAGCGGATGTGTGGCGACATTACTTTTCATGATGAGGAGCATGAA 692
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 567 ccctgcggaag 577
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 693 TCCAGCGGAAG 703
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 3

BG295496

LOCUS

DEFINITION 602392826F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4504638 5',
mRNA sequence.

ACCESSION BG295496

VERSION BG295496.1 GI:13057189

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NIH-MGC <http://mgc.nci.nih.gov/>. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999). Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: The Cepko Laboratory CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Plate: LLAM10377 row: c column: 07 High quality sequence stop: 753.

REFERENCE 1 (bases 1 to 1131)
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

1..1131
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4504638"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
261 a 335 c 375 g 160 t

BASE COUNT
ORIGIN

Query Match 63.3%; Score 438.4; DB 10; Length 1131;
Best Local Similarity 91.9%; Pred. No. 2e-75;
Matches 496; Conservative 0; Mismatches 41; Indels 3; Gaps 3;

QY 41 ggcgcggcagggccggcgagcccgggagtcggcgctgctggcgaggtg-gctg 99
DB 1 GACCAGCTGCTGGGCGCGGGCGAGCCCCGGGAGTCGGCGCTGCTGGACGGGTGCTG 60
QY 100 cagcg-gggcgtggccggggggcgggcgggcggtgagcggggctgcagggcccggt 158
DB 61 CAGCGTGGCGGTGGCGGGGGCGCGGTGGCGGGAGCGGGCGCTATCAGCCCCCTGT 120
QY 159 acgacaagatccgactcgcggccggactacagcg-gtgcggctgagccactgtca 217
DB 121 ACGGTGGATCCGAGTCCGCGCGCGGAATACGAAGCAGCGTCCGCGGTGAGCCACTGTCA 180
QY 218 ccacgcacatggtggcagcgccgtggcaggatccctggagcactgctgtatgtaaccca 277
DB 181 CCACGCACATGTGGGGCGCGCGTGGCAGGATCCTGGAGCATTCGNGATGTACCCGA 240
QY 278 tgcagtgcgtcaagaccggatgcagatctcacagctgaccagctgcccgcctatcgca 337
DB 241 TCAGTGCCTCAAGACCCGGATGCAGAGCTACAGCCTACCCAGCGCCGCTATCGGA 300
QY 338 atggtgtgagggccctctgagagattataagaacgagggcctatgagggccatgaggg 397
DB 301 ACGTGTGGAGGCTCTCTGGGAATCATGAGGACAGAGGGCCCTGTGGAGGCCCATGCGGG 360
QY 398 ggtgaacgtcacagaacagcgcgagggccctgcccacgccccttatttgcgtgtaag 457
DB 361 GGCTGAACGTCACAGCAACAGCGCGGGGCCCTGCCACGCCCTCTATTTTGCTGTACG 420
QY 458 aaaagttaaaagacattgagtgtgtaataccaccctgggggcaatgcccattattgcca 517
DB 421 AAAAGTTAAAAAGACATTGAGTGCAGTAATCACCCAGGGGGCAATAGCCATATTGCCA 480
QY 518 atgggtgcggcggtgtgtggcaacattacttcatgatacagccatgaacctgcggaag 577
DB 481 ATGGTGCACCCGGATGTGTGGGAGACTTACTTTCATGATGCAGCCATGAATCCAGCGGAAG 540

RESULT 4
BI541703

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
COMMENT

FEATURES
source

1..573
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 1BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
121 a 166 c 176 g 110 t

BASE COUNT
ORIGIN

Query Match 60.5%; Score 419.4; DB 10; Length 573;
Best Local Similarity 95.4%; Pred. No. 9.4e-72;
Matches 432; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 125 gcgcggggagcgccggcgctgcagggcccggtacgacaagatccggactccggccgg 184
DB 1 GCGGGGGGAGCGCGGGCGCTCAGGCCCTGTAGCAGGATCCGGACTCCAGGCCGG 60
QY 185 actacgagcgctgcggcgctggagccactgtcacacgcacatggtggcagggcgctgg 244
DB 61 ACTACGAGCGCTGCGGGCTGGAGTACTGTCCACGCACATGTTGGGGGGCGCGTGG 120
QY 245 caggatcctggagcactgctgatacccatgactgctcaagaccggatgcaga 304
DB 121 CAGGGATCTTGAGGACTGCGGTGATGACCCGCTGACTGCGTCAAGCCCGGATGCAG 180
QY 305 gtctacagcctgacccagctgccgctatcgcaatgtgttgagggcctcttgaggatta 364
DB 181 GCCTACAGCGGATCCAGCGCCCGCTATCGCAATGTGTGGAGGCCCTCTTGAGGATTA 240
QY 365 taagaacggagggcctatggagggcccatgaggggctgaactcagcaacagcgcgag 424

Note: this is a NIH MGC Library."

BASE COUNT	154	a	278	c		348	g		187	t
ORIGIN										
Query Match	57.5%	Score	398.2	:	DB 10;	Length	967	:		
Best Local Similarity	89.88	Pred.	No.	1.3e-67	Mismatches	48	Indels	6	Gaps	4:
Matches	473	Conservative	0							
OY	1	atggagttagagggcggttgccttgccggcgttcacgaaggccgccgcccggccccggg	60							
Dd	111	ATGAGATTGGAGGGCGAGTGCACGCCGTCTCCTCCGTCAGCAAGCATTTTGCTGCTG	170							
OY	61	cgaagaccggcgggactcggcgctgctcaaacggtttgctgc---agcggcgctgaggcccg	117							
Dd	171	CAGAACCCCCGGGAGTCCGGCGCTCTCTCACGGGTGGTCTTAGCGGTGGCGTGGCCCTGG	230							
OY	118	ggggcgccggcgggggagccggggcctgcagggcccccggttac-gacaagaatccggacct	176							
Dd	231	GGSGCGCGCGCGAGCGGGSCCTATCAGCCCCCTGTACTGGCTGGATCCGGAGTC	290							
OY	177	cggccggactaaagggcgttcggcggtcgagccaattcacccacgcacatggtggcagg	236							
Dd	291	CGCGCCGAATACAAAGCGTTCGCGGTGGAGCCACTGTACCACGCACAT-GTGGCGGG	349							
OY	237	cgcgtgcagggatcctgagcaactcgtgatatacccatcgaatcgtcgaagaccgc	296							
Dd	350	CGCGTGGCAGGGATCTGGAGCATTCGGTGTATCCCAGATCGACTGCTCAAGACCCG	409							
OY	297	gatcagagctcacagcctgacccagctgcgcgctatcgaatgttgtggaggccctctg	356							
Dd	410	GATGCAGAGCTTACAGCTTGACCCAGCGCCCGCTATCGGAACGTGTGGAGGCTCTCTG	469							
OY	357	gaggtattaagaacggagggccttatggagcccatgagggggtgtaacgtcacagaac	416							
Dd	470	GAGAATAATGAGACAGAGGGCCCTGTGGAGGCCATCGGGGGCTGAACGCTCAGCAAAC	529							
OY	417	aggcgagggcctgcgcacccctttatttgctcgtacgaaaagttaaaaagacatt	476							
Dd	530	AGCGCGGGCGCTGCCACGCCCTCTATTTCCTGCTGCTACGAAAAGTTAAAAAAGACATT	589							
OY	477	gagtgtgtaatcacocctgggggc-aatagccatatggccaatggt	522							
Dd	590	GAGTGAGCTTAATCACCGAGGGGCAATAGCCATATTGCCAATGGT	636							
RESULT	7									
LOCUS	BG964218									
DEFINITION	602828936F1 NCI_CGAP_Co24 Mus musculus cdna clone IMAGE:4903815 5'	645 bp	mRNA	linear	EST 12-JUN-2001					
ACCESSION	BN964218									
VERSION	BN964218.1	GI:14351855								
KEYWORDS	EST.									
SOURCE	house mouse.									
ORGANISM	Mus musculus									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
TITLE	1 (bases 1 to 645)									
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)									
COMMENT	Contact: Robert Strauszberg, Ph.D. Email: cgapbs-r@mail.nih.gov CDNA Library Preparation: Jeffrey E. Green, M.D. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10988 row: p column: 24 High quality sequence start:2									


```
RESULT 10
AV704087
LOCUS      681 bp      mRNA      linear      EST 09-OCT-2000
DEFINITION AV704087 ADB Homo sapiens cDNA clone ADBAF05 5', mRNA sequence.
ACCESSION  AV704087
VERSION     AV704087.1 GI:10721407
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 681)
AUTHORS   Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao
           H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
           Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu
           G., Hu, R., Chen, J., Chen, Z. and Han, Z.
TITLE     Homo sapiens cDNA ADB clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zeguang Han
           Chinese National Human Genome Center at Shanghai
           351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
           201203, P. R. China
           Tel: 86-21-50801919(ex.45)
           Fax: 86-21-50801922
           Email: hanzg@chgc.sh.cn
           This clone is available at CHGC in Shanghai.
FEATURES   source
            1..681
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="ADBAFE05"
            /clone_lib="ADB"
            /tissue_type="Adrenal gland"
            /dev_stage="adult"
            /lab_host="SOUR"
            /note="vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
            XhoI"
BASE COUNT 149 a 209 c 205 g 117 t 1 others
ORIGIN
Query Match      50.0%; Score 346.4; DB 9; Length 681;
Best Local Similarity 99.7%; Pred. No. 1.4e-57;
Matches 347; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 116 gggggccgcggcgggagccggggcctgcaggcccccgggtacgacaagatcggact 175
Db 1 GGGGGCCGCGGGGAGGCGGGGCTGCAGGCCCCCGGTACGCAAGATCCGGACT 60
Qy 176 ccggcccgactacgagcgcctgcggctggagccactgtcaccacgcacatggtggcag 235
Db 61 CCGGCCCGACTACGAGCGCTGCGGGTGGAGCCACTGTCAACACGCACATGTGGGAG 120
Qy 236 gcgcgtggcggagatcctggagcactgcgtgtatgtatcccccatgcactgcgtcaagacc 295
Db 121 GCGCGGAGCAGGATCCTGGGACACTGCGTGATGTACCCCATCGACTGCTCAAGACCC 180
Qy 296 ggatgcagagctcagcctgaccagctgcccagctgcccgtatcgcaatgtgtggagccctct 355
Db 181 GGATGCAGAGTCTACAGCTTACCCAGCTGCCCCGTATCGCAATGTGTGGAGGCCCTCT 240
Qy 356 ggagattataagcaacgagggcctatgagggcccatgagggcctgaacgtcacagcaa 415
Db 241 GGAGGATTATAGACGAGGGGCTATGGAGGCCCATGGAGGCCCATGGAGGGGCTGAACGTCAAGCAA 300
Qy 416 caggcgcaggcctgccacacccctttatttggctgctacgaaaaagt 463
Db 301 CAGGCGCAGGGCCTGCCACACGCCCTTTATTATTGGCTGCTACGAAAAGT 348
RESULT 11
BE913718
LOCUS      598 bp      mRNA      linear      EST 29-SEP-2000
DEFINITION BE913718 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5',
           mRNA sequence.
ACCESSION  BE913718
VERSION     BE913718.1 GI:10411618
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 598)
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE     NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Gilbert Smith, Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM9146 row: o column: 17
           High quality sequence stop: 593.
FEATURES   Location/Qualifiers
            1..598
            /organism="Mus musculus"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:3969256"
            /clone_lib="NCI_CGAP_Mam1"
            /tissue_type="tumor, biopsy sample"
            /dev_stage="10 months, virgin"
            /lab_host="DH10B"
            /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
            Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
            Library constructed by Life Technologies. Investigator
            providing samples: Gilbert Smith, NIH"
BASE COUNT 141 a 175 c 172 g 110 t
ORIGIN
Query Match      49.7%; Score 344.6; DB 10; Length 598;
Best Local Similarity 93.7%; Pred. No. 3e-57;
Matches 359; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 195 gctgccggctggagccactgtcacacgcacatggtggcaggcccggtggcaggatcct 254
Db 1 GCTGCCGCTGGAGCCACTGTCAACGCACATGCTGGCGGGCGCCGTGGCAGGGATCCT 60
Qy 255 ggagcactgcgtgatgtaccctacgcactgcgtcaaaccccgatcgagctctacagcc 314
Db 61 GGAGCTTTCGTGATGTACCCGATCGACTGCGTCAAGACCCGGATGCAGAGCCTACAGCC 120
Qy 315 tgaccagctcccgcctatcgcaatgtgttgaggccctctggaggattataaagacgga 374
Db 121 TGACCCAGCCCGCGCTATCGGAACGTGTGGAGGCTCTCTGGAGAAATCATGAGACAGA 180
Qy 375 gggcctatggaggccctatgaggggctgaacgctcacagcaacagcgccgagccctgccca 434
Db 181 GGGCCTGTGGAGGCCCATGCGGGGCTGAAGCTCACAGCAACAGCGCGGGCGCTGCCCA 240
Qy 435 cgccctttatttccctgcgtacgaaaaagttaaaaagacattgagtgatgtaaccaccc 494
Db 241 CGCCCTCTATTTTCCTGCTACGAAAAGTTAAAAAAGACATTTAGTGACGTAAATCCACC 300
Qy 495 tgggggcaatagccatattgccaatggtggccggcggtgtgtggcaacattactcatga 554
Db 301 AGGGGGCAATAGCCATATTGCCAATGGTGCAGCCCGATGTGTGGCAGATTACTTCTATGA 360
Qy 555 tgcagccatgaaccctgcggaag 577
||||||||||| |||||||||
```


Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES

source

```

1. .546
/organism="Gallus gallus"
/strain="Commercial broiler and Ottawa Res. Centre
strains 90 & 21"
/db_xref="taxon:9031"
/clone="pgm2n.pk008.12"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Brest, leg; Embryo(d19); post-hatch(1d,1,3,5,7,9
,11 weeks); growth plate(1d,7d,14d post-hatch)"
/lab_host="E. coli EMD10B"
/notes="vector: pcMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5' end"
```

BASE COUNT	90 a	183 c	197 g	70 t	6 others
ORIGIN					

Query Match 43.1%; Score 298.4; DB 10; Length 546;
Best Local Similarity 77.0%; Pred. No. 2.8e-48;
Matches 375; Conservative 0; Mismatches 109; Indels 3; Gaps 1;

[illegible]

Search completed: September 28, 2002, 06:22:41
Job time: 9182 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:55:42 ; Search time 5287.54 Seconds
(without alignments)
3158.251 Million cell updates/sec

Title: US-09-870-113-7
Perfect score: 798
Sequence: 1 atgcagagtctacagctga.....agtggaggctggcaagtga 798

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	798	100.0	1322	6	AX061229	Sequence
2	798	100.0	1448	9	AF327402	Homo sapi
3	798	100.0	1448	17	HS303077	Homo sapi
4	576	72.2	1889	9	AF327403	Homo sapi
5	576	72.2	1889	17	HS303078	Homo sapi
6	520	65.2	1244	9	AF267854	Homo sapi
7	520	65.2	123160	9	AL353719	Human DNA
8	518.4	65.0	1588	9	AK056782	Homo sapi
9	504	63.2	1072	10	BC023172	Mus muscu
10	467.2	58.5	221062	2	AC096351	Rattus no
11	417.8	52.4	483	6	AX306258	Sequence
12	392	49.1	401	6	AX071440	Sequence
13	387.8	48.6	4112	10	AF288621	Mus muscu
14	365.6	45.8	1429	9	AF223466	Homo sapi
15	281	35.2	5869	10	AF361699	Mus muscu
16	276	34.6	1410	9	AF155660	Homo sapi
17	276	34.6	1410	9	AF155660	Homo sapi
18	251.2	31.5	191702	2	AC108878	Mus muscu
19	250	31.3	174445	9	AC051642	Homo sapi
20	250	31.3	182504	2	AC022597	Homo sapi
21	248.4	31.1	108765	2	AF216674	Homo sapi
22	166.8	20.9	818	6	AX012190	Sequence
23	155.6	19.5	434	11	G28167	human STS S
24	153.6	19.2	338	6	AX203265	Sequence
25	150.8	18.9	34305	3	CEW02B12	Caenorhabd
26	141.4	17.7	3037	3	AY060268	Drosophil
27	134.6	16.9	1281	3	OYU45998	Onchocerca
28	125.2	15.7	163205	2	AC093368	Rattus no
29	113.4	14.2	599	6	AX198694	Sequence
30	113.4	14.2	599	6	AX209226	Sequence
31	113.2	14.2	2878	3	AF217402	Drosophil
32	105	13.2	108765	2	AF216674	Homo sapi
33	104	13.0	1243	3	OGU45997	Onchocerca
34	103.4	13.0	858	9	BC015013	Homo sapi
35	91.4	11.5	391	11	G22871	human STS W
36	86	10.8	2096	8	SCMR33G	Yeast nucle
37	86	10.8	2496	8	SCYJL133W	S.cerevisia
38	86	10.8	40724	8	SCX731	S.cerevisia
39	84.8	10.6	500	6	AX336750	Sequence
40	84.8	10.6	657	6	AX011696	Sequence
41	84.6	10.6	1864	8	SCYKR052C	Sequence
42	84.6	10.6	2601	8	SCMR34	S.cerevisia
43	82.2	10.3	76299	2	AC019585	Yeast nucle
44	82.2	10.3	172363	3	AC009388	Drosophil
45	82.2	10.3	203557	3	AE003763	Drosophil

ALIGNMENTS

RESULT 1	AX061229	AX061229	Sequence	76 from Patent WO0078953.	DNA	linear	PAT 22-JAN-2001
LOCUS	AX061229	Sequence	76 from Patent WO0078953.	1322 bp	DNA	linear	PAT 22-JAN-2001
DEFINITION	AX061229	Sequence	76 from Patent WO0078953.	1322 bp	DNA	linear	PAT 22-JAN-2001
ACCESSION	AX061229	Sequence	76 from Patent WO0078953.	1322 bp	DNA	linear	PAT 22-JAN-2001
VERSION	AX061229.1	GI:12406365					
KEYWORDS							
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	1 (bases 1 to 1322) Lal, P., Yang, J., Yue, H., Hillman, J.L., Tang, Y.T., Bandman, O., Burford, N., Baughn, M.R., Azimzai, Y., Lu, D.A., Au-Young, J. and Patterson, C.						
TITLE	Human transport proteins						
JOURNAL	Patent: WO 0078953-A 76 28-DEC-2000;						
FEATURES	Incyte Genomics, Inc. (US)						
Source	Location/Qualifiers						
	1..1322						
	/organism="Homo sapiens"						
	/db_xref="taxon:9606"						
BASE COUNT	328 a	376 c	351 g	267 t			

ORIGIN		SOURCE	human...
		ORGANISM	Homo sapiens
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
		REFERENCE	1 (bases 1 to 1448)
		AUTHORS	Li,F.Y., Nikali,K., Grogan,J., Leibiger,I., Leibiger,B., Schweyen,R., Larsson,C. and Suomalainen,A.
		TITLE	Characterization of a novel human putative mitochondrial transporter homologous to the yeast mitochondrial RNA splicing proteins 3 and 4
		JOURNAL	FEBS Lett. 494 (1-2), 79-84 (2001)
		MEDLINE	21195335
		PUBMED	11297739
		REFERENCE	2 (bases 1 to 1448)
		AUTHORS	Li,F., Nikali,K., Grogan,J., Leibiger,I., Leibiger,B., Schweyen,R., Larsson,C. and Suomalainen,A.
		TITLE	Direct Submission
		JOURNAL	Submitted (08-DEC-2000) Human Molecular Genetics, National Public Health Institute, Mannerheimintie 166, Helsinki 00300, Finland
		FEATURES	Location/Qualifiers
		source	1..1448
			/organism="Homo sapiens"
			/db_xref="taxon:9606"
			/chromosome="10"
			/map="10q24"
			30..1124
			/note="similar to mitochondrial RNA splicing protein 3/4; alternatively spliced"
			/codon_start=1
			/product="putative mitochondrial solute carrier splice variant"
			/protein_id="AAK49519.1"
			/db_xref="GI:13926047"
			/translation="MELEGRAGVAGVAGGPAAGPCRPSPGESALLDGNLQRVGRGAGGG
			EGACRPVRRQDPDSDGPDYEALPAGATVTTHMVAGVAGLEHCVMPIDCVKTRMOS
			LQDPDAARYRNLEALWRIIRTEGLMRPMGLNVTATGAPAHALFYACYEKLKKTLS
			DVTHPGNSHIANGAACVATLLHDAAMPNPAEVVKORMQNSPYHRVTPCVRVWON
			EGAGAFVSVTTLTMNVPOAHEMTYFLOEHPNRPORRYNPSHVLSCACAGAVAA
			AATPLDVCKTLNTOESLANSHTGHITGMASAFYVQGVGTAYFRGVQARVYI
			QIPSTALIAVSYIEFFLIITKROEENRAGK"
		BASE COUNT	322 a 408 c 435 g 283 t
		ORIGIN	
		Query Match	100.0%; Score 798; DB 9; Length 1448;
		Best Local Similarity	100.0%; Pred. NO. 7.2e-240;
		Matches 798; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	atgcagagctctacagcctgacccagctgacccagctgcccgcctatgcgaatgtgttgaggccctctgg	60
Db	327	ATGCAGAGCTCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG	386
Qy	61	aggattataagaacgaggggcctatgaggcccatgaggcccatgaggggcctgaacgtcacagcaaca	120
Db	387	AGGATTATAAGAACGGAGGGCCTATGAGGGCCCATGAGGGGCTGAACGTACACAGCAACA	446
Qy	121	ggcgagggcctgccacgccccttttttgcctgcctacaaaaagttaaaaaacattg	180
Db	447	GGCGAGGGCCTGCCACGGCCCTTTATTTGCCCTGCTACGAAAAGTTAAAAAGACATTG	506
Qy	181	agtatgtaatccaccctgggggcaatagccatatgtccaatggtgcggcggtgtgtg	240
Db	507	AGTGATGTAATCCACCTGGGGGCAATAGCCATATGCCAATGTGCGCGGGTGTGTG	566
Qy	241	gcaacattactctatgatgcagcccatgaacccctcggaagtgttcaagcgagagatgcag	300
Db	567	GCAACATTACTTCTATGATGCAGCCATGAACCTCGCGAAGTGTGTCAGAGCAGAGGATGCAG	626
Qy	301	atgtacaactcacatcacaccgggtgcacagactgtgtacggcgagtggtgcaaatgaa	360
Db	627	ATGTACAACCTCACCATACCGGGGTGACAGACTGTGTACGGGGCAGTGTGGCAAAATGAA	686
Qy	361	ggggccggggcctttttaccgcagctcacaccacccagctgaccatgaacgttcctcttccaa	420

RESULT 2
AF327402
LOCUS
DEFINITION
Homo sapiens putative mitochondrial solute carrier splice variant
mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.
ACCESSION
AF327402
VERSION
AF327402.1
KEYWORDS
GI:13926046

|||||
Db 687 GGGGGCGGGCCCTTTACCGCAGCTACACCCAGCTGACCATGAACGTTCTCTTCCAA 746
QY 421 gccattcattcatgacctatgaattcctgagagacatttaacccccagagaggtac 480
Db 747 GCCATTCACTTTCATGACCTATGAATCTCTGCGAGGACCTTTAACCCCGAGAGCGGTAC 806
QY 481 aacccaagctcccaagctctctctgagcttgcgagagctgtagctgcccagccaca 540
Db 807 AACCCAAAGCTCCACGCTCTCTGGAGCTTGGCAGGAGCTGTAGCTGCCAGCCACA 866
QY 541 accccactggagctttgcaaaaactgctcaaacccccagagagctcttggcttgaactca 600
Db 867 ACCCCACTGGAGCTTTGCAAAACACTGCTCAACACCCAGGAGCTCTTGGCTTTGAACCTCA 926
QY 601 cacattacagacatcacagcgatgctgtagtgccttcagagacgttatcatcaagtagt 660
Db 927 CACATTACAGGACATATCACAGCATGCTGTAGTGCTTCAGGACGGTATATCAAGTAGGT 986
QY 661 ggggtgacgcctattccgaggggtgcagccagagtaatttacagatccccctccaca 720
Db 987 GGGGTGACCGCCTATTTCGAGGGGTGCAGGCCAGAGTAATTTACAGATCCCTCCACA 1046
QY 721 gccatcgatgctgctgtatgatttctcaaatcactaaacaggaagcaagaag 780
Db 1047 GCCATCGCATGCTGCTGTATGAGTCTTCAAAATACCTAATCACTAAAAGGCAAGAAG 1106
QY 781 tggagggtggcgaagtga 798
Db 1107 TGGAGGGCTGGCAAGTGA 1124
RESULT 3
HSA303077
ID HSA303077 standard; RNA; HUM; 1448 BP.
XX
AC AJ303077;
XX
SV AJ303077.1
XX
DT 02-FEB-2001 (Rel. 66, Created)
DT 26-APR-2001 (Rel. 67, Last updated, Version 2)
DE Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMRS3/4
DE gene), 1448 bp cDNA splice variant
XX
KW HMRS3/4 gene; mitochondrial RNA splicing protein 3/4.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]
RP 1-1448
RA Nikali K.;
RT ;
RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.
RL Nikali K., Human Molecular Genetics, National Public Health Institute,
RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.
XX
RN [2]
RA Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
RA Larsson C., Suomalainen A.;
RT "Characterization of a novel human putative mitochondrial transporter
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";
RL FEBS Lett. 494:79-84(2001).
XX
FH Key Location/Qualifiers
FH
FH 1. .1448
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT 30. .1124
FT CDS

/gene="HMRS3/4"
/product="mitochondrial RNA splicing protein 3/4"
/function="putative mitochondrial solute carrier"
/protein_id="CAC21996.1"
/translation="MELEGRAGVAGGPAAGPGRSPGESALLDGLWLRGVGRGAGGE
AGACRPVRQDPDSDPYEALPAGATVTHMVAGAVAGILEHCYVMPYIDCVKTRMSIQ
PDPAARYNLEALWRIITLGLRPMRLNVTATGAGPAHALYFACYEKLKLTLSDYI
HPGNSHIANGAAGCAVATLLHDAAMPNPAEVVKQRMOMVNSPYHRVTDGCRVWONEGAG
AFYRSYTTQLTMNYPFOAIHEPMTYEFLEHENPQRRYNSPSSHVLSGACAGAAAAATTP
LDVCTLLENTOSLALNSHITGTMASAFRTVYQGVGTAYTFRGVARVYIQLPSTA
IAWSYEFFKYLITRKQBEWRAG"
SQ Sequence 1448 BP; 322 A; 408 C; 435 G; 283 T; 0 other;
Query Match 100.0%; Score 798; DB 17; Length 1448;
Best Local Similarity 100.0%; Pred. No. 7.2e-240;
Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgcagagctctacagcctgaccagctgccgctatcgcaatgctgttgaggccctctg 60
Db 327 ATGCAGAGCTTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 386
QY 61 aggattataagaacagggggcctatgagggcccatgaggggctgaacgtcacacaaca 120
Db 387 AGGATTATAAGAAACGAGGGCCTATGGAGGCCCATGAGGGGCTGAACCTCACGCAACA 446
QY 121 ggcgagggcctgcccagcccttatttgcctgctacgaaaagttaaaaagacattg 180
Db 447 GCGCAGGGCCTGCCACGCGCCTTTATTGTGCTGCTACGAAAGCTTAAAAAGACATTG 506
QY 181 agtgaatgaatccacctgggggcaatagccatattgccaatggtgcggccgggtgtg 240
Db 507 AGTGATGTAATCCACCTGGGGGCAATAGCCATATGCCAATGTGGTGGCGCCGGTGTGG 566
QY 241 gcaacattacttcattgatgcagccatgaacctgcggaagtgtgaagcagaggtatgcag 300
Db 567 GCAACATTACTTCATGATGCAGCCATGAACCTCGGGAAGTGTCAAGCAGAGGATGCAG 626
QY 301 atgtacaactcaccataccaccgggtgacagactgtgtacggggcaggtggtgcaaatgaa 360
Db 627 ATGTACAATCACCATACCCAGCGGTGACAGACTGTGTACGGGCGAGTGTGCGCAAAATGAA 686
QY 361 ggggcggggcctttaccgcagctcacaccaccagctgacctgaacgttctcttccaa 420
Db 687 GGGGCGGGGCGCTTTACCGCAGCTACACCACCCAGCTGACCATGACGTTCTTCCAA 746
QY 421 gccattcattcatgacctatgaattcctgcaggagcactttaacccccagagagcgttac 480
Db 747 GCCATTCACTTTCATGACCTATGAATTCCTGCGAGGACCTTTAACCCCGAGAGCGGTAC 806
QY 481 aacccaagctcccaagctctctctgagcttgcgagagctgtagctgcccagccaca 540
Db 807 AACCCAAAGCTCCACGCTCTCTGTGAGCTTGGCAGGAGCTGTAGCTGCCAGCCACA 866
QY 541 accccactggagctttgcaaaaactgctcaaacccccagagagctcttggcttgaactca 600
Db 867 ACCCCACTGGAGCTTTGCAAAACACTGCTCAACACCCAGGAGCTCTTGGCTTTGAACCTCA 926
QY 601 cacattacagacatcacagggcatggtgtagtgccttcagagcgggtatatcaagtagt 660
Db 927 CACATTACAGGACATATCACAGGCATGGCTAGTGCCTTCAGGACGGTATATCAAGTAGGT 986
QY 661 ggggtgacgcctattccgaggggtgcagccagagtaatttacagatccccctccaca 720
Db 987 GGGGTGACCGCCTATTTCGAGGGGTGCAGGCCAGAGTAATTTACAGATCCCTCCACA 1046
QY 721 gccatcgatgctgctgtatgatttctcaaatcactaaacaggaagcaagaag 780
Db 1047 GCCATCGCATGCTGCTGTATGAGTCTTCAAAATACCTAATCACTAAAAGGCAAGAAG 1106
QY 781 tggagggtggcgaagtga 798
Db 1107 TGGAGGGCTGGCAAGTGA 1124

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

<http://www.sanger.ac.uk/Projects/C.elegans/wormpep> This sequence

was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr10>

RP11-85A1 is from the library RPI1-11.1 constructed by the group of Pieter de Jong. For further details see

<http://www.ncbi.nlm.nih.gov/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone

RP11-85A1. It may be shorter because we sequence overlapping

sections only once, except for a short overlap.

The true left end of clone RP11-85A1 is at 1 in this sequence. The

true left end of clone RP11-483F11 is at 123061 in this sequence.

The true right end of clone RP11-129J12 is at 51589 in this

sequence.

```

FEATURES             Location/Qualifiers
     source            1..123160
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="10"
                        /clone="RP11-85A1"
                        /clone_lib="RPI1-11.1"
misc_feature          complement(7066..7118)
                        /note="Single clone region. Sequence from reads from a
                        short insert library derived from a single pUC clone.
                        Restriction digest data confirm the assembly."
     unsure            52162
     unsure            52167
BASE COUNT           33221 a 28084 c 27196 g 34659 t
ORIGIN

Query Match          65.2%; Score 520; DB 9; Length 123160;
Best Local Similarity 100.0%; Pred. No. 3.5e-152;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  279 agtggtaagcagagagatgcagatgtacaactaccataccacccgggtgacagactgtgt 338
Db  82990 AGTGGTCAAGCAGAGAGATGTCAGATGTACAAGTCACTACCATCACCACCGGGTGCAGACACTGTGT 82931

Qy  339 acgggcagtggtgcaaatgaaggcgccgtttttaccgcagctacacacccagct 398
Db  82930 ACGGGCAGTGTGGCAAAATGAAGGGGCGGGCGCTTTTACCAGCAGCTACACACCCAGCT 82871

Qy  399 gaccatgaacgttcccttcccaagccattcacttcacgtacacattccttcctgcaggagca 458
Db  82870 GACCATGAACGTTCTTCTTCCAAAGCATTCACTTTCATGACCTATGAATTCCTGCAGGAGCA 82811

Qy  459 ctttaacccccagagagcgtgtacaaccaagctccacgctcctctctgtgagcttgcgagg 518
Db  82810 CTTTAAACCCCCAGAGAGCGGTACAACCAAGCTCCACAGCTCCTCTCTGTGAGCTTGCAGG 82751

Qy  519 agctgtagctccgcagccacacccactgagctgtttgtgaaaaactgctcaacaccca 578
Db  82750 AGCTGTAGTCTCCGACGCCAACCCCACTGGAGTTTGTGAAAAACACTGCTCAACACCCA 82691

```

```

Qy  579 ggagtccttggttgaaactcacacattacagagacatcatcacagcagcatggctagtcctt 638
Db  82690 GGAGTCCTTGCTTGAACCTCACACATTACAGGACATATCACAGCATGGCTAGTGCTT 82631

Qy  639 caggacggtatatacaagttagtggtgggtgacgcctatttccgaggggtgcagggcagagt 698
Db  82630 CAGGACGGTATATCAAGTAGTGGGTGACGCCCTATTTCGAGGGGTGCAGGCAGAGT 82571

Qy  699 aattaccagatccctccacagcagcatgcagctgctgtgtatgatgttcttcaataacct 758
Db  82570 AATTACCAGATCCCTCCACAGCATGCGCATGCTGTGTATGAGTTCTTCAATAACCT 82511

Qy  759 aatcactaaaggcaagaagagtgagggtgcgtgcgaagtga 798
Db  82510 AATCATAAAGGCCAAGAGAGTGGAGGGCTGGCAAGTGA 82471

```

RESULT 8

AK056782

LOCUS

DEFINITION Homo sapiens cDNA FLJ32220 fis, clone PLACE6003850, moderately similar to Homo sapiens mitochondrial solute carrier mRNA.

ACCESSION AK056782

VERSION AK056782.1 GI:16552284

KEYWORDS Oligo capping; fis (full insert sequence).

SOURCE Homo sapiens placenta cDNA to mRNA, clone_lib:PLACE6

clone:PLACE6003850.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,

Ono, Y., Hotuta, T., Hiraoka, S., Murakawa, K., Takiguchi, S.,

Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,

Yamashita, H., Chiba, Y., Sugiyama, T., Irie, K., Otsuki, T., Sato, H.,

Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y.,

Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y.,

Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A.,

Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,

Nagabari, K., Masubo, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 1588)

Isogai, T., Otsuki, T. and Sugiyama, T.

Direct Submission

Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,

Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3951, Fax: 81-438-52-3952)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: RAB and

HRI.

FEATURES

source

1..1588

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="PLACE6003850"

/tissue_type="placenta"

/clone_lib="PLACE6"

/notes="cloning vector: pME18SFL3"

BASE COUNT 385 a 405 c 375 g 423 t

ORIGIN

Query Match 65.0%; Score 518.4; DB 9; Length 1588;

Best Local Similarity 99.8%; Pred. No. 8.7e-152;

Matches 519; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 279 agtggtaagcagagagatgcagatgtacaactaccataccacccgggtgacagactgtgt 338

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

. Location/Qualifiers

1. .1072

/organism="Mus musculus"

```
/db_xref="taxon:10090"
```

```
/clone="MGC:37028 IMAGE:4949779"
```

/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma, 5 month old virgin mouse."

```
/clone_lib="NCI_CGAP_Mam
```

```
/lab host="DH10B"
```

```
/tab_node = D110D
/note="Vector: pCMV-SPORT6"
```

188. 721

```
100: ./Z1
/codon start=1
```

```
/product="Similar to putative mitochondrial solute
```

carrier"

carrier
protocol

```

/protein_id="AAH231/2.1"
/db_xref="GI:18606248"
/translation="MNDAEIVKQPMQMVNSDYDPIVTDGVDAITQNEFCACAEVDSVTTQ

```

CTABSCTCIIH="MNPAAEVVKQRMQMINSPYHRVI-DCVRVWQNEGAGAFYRSY-TIQ
ITMNVBEQAIHEMTYEEIOEHENBOPBYNPDESHVICCACAQADAAATRI-DVCCTII

DIMNVFFQAIHFMI IEF LQEHFNFQKRI
 NTOESI.AINSNTTGHITGMA SAEFTVY

Query Match

Query Match	63.2%;	Score 304;	DB 10;	Length 1072;
Best Local Similarity	92.28;	Pred No	2	98-147.

Best Local Similarity 92.28; Pred. NO. 2.9e-147;
Matches 531. Conservative 0. Mismatches 15. Indels 0.

y 223 ggtgcgccgggtgtgtggcaacattacttcatgatgcagccatgaaccctgcggaagtg 282

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

283 gtcaagcagaggatgcagatgtacaactcaccacgggtgacagactgtgtacgg 342

.....

y 343 gcagtgtggcaaaatgaaggqggccqggccctttaccqcaqctacaccaccccaqctqacc 402

D 266 GCAGTGTGGCAAAATGAAGGGGCCGGGGCCCTTTTACCGCAGCTACACGACCCAGCTGACC 325

[illegible]

b 326 ATGAATGTCCCTTCCAAGCCATTCACTTCATGACCTATGAGTTCCTGCAAGAGCACTT 385

[illegible]

386 AACCCCAGAGACGGTACAACCCCAGCTCCCCATGTGCTCTGTGGGGCTGCCGAGGAGCA 445

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

b 446 GTAGCTGACGCCGCCACAACCCCACTGGACGTTTGTAAACACTGCTCAACACCCAGGAA 505

583 tccttggcttgaactcacacattacaggacatatcacaggcatggctagtgcccttcagg 642

[illegible]

643 acqqtatatcaadttaggttqqqqtqaccqccctatttccqaqqqqtqcaqqccagaqtaatt 702

F

703 taccagatccctccacagccatggtctgtgtatgagttcttcaataacctaac 762

RESULT 10

```
AC096351/c
LOCUS
DEFINITION
Rattus norvegicus chromosome Rf1 clone CH230-24M6, HTG 20-DEC-2001
SEQUENCE, 33 unordered pieces.
ACCESSION
AC096351
VERSION
GI:17944054
KEYWORDS
HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 221062)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Albrooks, S.L., Amaral, H.C., Are, J.R., Banks, T., Barbarella, J.,
Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowle, S., Briveau, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burke, P., Burket, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, X., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,
Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
Joudah, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Loulsegh, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Mar, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Massey, E., McLeod, M.P., Meador, M.,
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,
Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Prins, E., Pu, L.L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,
Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q40
```

```
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 33 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved
  1
  32851: contig of 32851 bp in length
  32852 32951: gap of unknown length
  32952 49984: contig of 16033 bp in length
  49985 49984: gap of unknown length
  63815: contig of 14731 bp in length
  63915: gap of unknown length
  63916 76005: contig of 12090 bp in length
  76006 76105: gap of unknown length
  88842 88841: gap of unknown length
  88842 100591: contig of 11650 bp in length
  100592 100691: gap of unknown length
  100692 111702: contig of 11011 bp in length
  111703 111802: gap of unknown length
  111803 122664: contig of 10862 bp in length
  122665 122764: gap of unknown length
  122765 129711: contig of 6947 bp in length
  129712 129811: gap of unknown length
  129812 13247: contig of 7436 bp in length
  13248 137347: gap of unknown length
  137348 144676: contig of 7329 bp in length
  144677 144776: gap of unknown length
  144777 151224: contig of 7348 bp in length
  151225 152224: gap of unknown length
  152225 158878: contig of 6654 bp in length
  158879 158978: gap of unknown length
  158979 163928: contig of 4950 bp in length
  163929 164028: gap of unknown length
  164029 169353: contig of 5325 bp in length
  169354 169453: gap of unknown length
  169454 175590: contig of 6137 bp in length
  175591 175690: gap of unknown length
  175691 180475: contig of 4785 bp in length
  180476 180575: gap of unknown length
  180576 184657: contig of 4082 bp in length
  184658 184757: gap of unknown length
  184758 190332: contig of 5575 bp in length
  190333 190432: gap of unknown length
  190433 194134: contig of 3702 bp in length
  194135 194234: gap of unknown length
  194235 196592: contig of 2358 bp in length
  196593 196692: gap of unknown length
  200530 200530: contig of 3838 bp in length
  200531 200630: gap of unknown length
  200631 202809: contig of 2179 bp in length
  202810 202909: gap of unknown length
  202910 205034: contig of 2125 bp in length
  205035 205134: gap of unknown length
  205135 207176: contig of 2042 bp in length
  207177 207276: gap of unknown length
  207277 208845: contig of 1569 bp in length
  208846 208945: gap of unknown length
  211271 211271: contig of 2326 bp in length
  211272 211371: gap of unknown length
  213302 213302: contig of 1931 bp in length
  213402 213402: gap of unknown length
  213403 215484: contig of 2082 bp in length
```

```
* 215485 21584: gap of unknown length
* 21585 21699: contig of 1385 bp in length
* 21699 21706: gap of unknown length
* 21706 21835: contig of 1289 bp in length
* 21835 21858: gap of unknown length
* 21858 21859: gap of unknown length
* 21859 21886: contig of 1428 bp in length
* 21886 21988: gap of unknown length
* 21988 22106: contig of 1076 bp in length.
* 21987 22106: contig of 1076 bp in length.
FEATURES
  Location/Qualifiers
    1..22106
      /organism="Rattus norvegicus"
      /db_xref="taxon:10116"
      /chromosome="R1"
      /clone="CH230-24M6"
BASE COUNT 58389 a 50577 c 49732 g 58877 t 3487 others
ORIGIN

Query Match      58.5%; Score 467.2; DB 2; Length 221062;
Best Local Similarity 93.7%; Pred. NO. 1.6e-135;
Matches 487; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 279 agtgtcaagcagagatgcagatgacactcaccataccaccgggtgacagactgtgt 338
DB 99390 AGTGTCAAGCAGCGATGCATGTATCAACTACCGTACCACCGGGTGACAGACTGTGT 99331

QY 339 acgggcagtggtgcaaaatgaaggcgccgtttaccgcagctacaccaccagct 398
DB 99330 TCGGCGAGTGTGGCAAAATGAAGGGCGCGGGCCCTTTTACCGCAGCTACACGACCAGCT 99271

QY 399 gaccatgaacgtctcttcacagccattacattcatgacattacattcctcgcaggaca 458
DB 99270 GACCATGAATGTTCCTTCCAAAGCCATTACCTTCATGACCTATGAGTTCCCTCGAGGACA 99211

QY 459 ctttaaccacagacaggtacaccccaagctcccaagctctctctgtgagcttgcagg 518
DB 99210 CTTTAAATCCCCAGACGGGTACACCCAGCTCCCATGTGCTGTGGAGCGTGGCAGG 99151

QY 519 agctgtagtcgcgcagccacacacccactggacgttttgcaaaacactgctcaacaccca 578
DB 99150 AGCTGTAGTCGCGCGCCACACCCACTGGACGTTTGCAAAACACTGCTCAACACCCA 99091

QY 579 ggaagctctgtgttgaaactcaacattacagggacatcacagggcatggttagtgcctt 638
DB 99090 GGAGTCCCTGGCCTTGAGTCAAAATTAACAGGACACATCACAGGCATGGCTTAATGCCTT 99031

QY 639 caggacgtatatacaagtagtggtgacccgtatttcgaggggtgcaggccagagt 698
DB 99030 CAGGACGGTCTATCAAGTAGGGGGGTGACTGCTTACTTCCGAGGGGTACAGGCCAGAGT 98971

QY 699 aattaccagatccctccacagccatgcgcatggctgtgtatgagttcttcaaatacct 758
DB 98970 AATTTACCAGATCCCTCCACAGCCATCGCATGGTCTGTGTATGAATCTTCAATAACCT 98911

QY 759 aatcaactaaaaggcaagagagtgaggggctggcgaagtga 798
DB 98910 AATCAAGAGCGGCAAGAGATGAGGGCGAGGCAAGTGA 98871

RESULT 11
AX306258 AX306258 483 bp DNA linear PAT 11-DEC-2001
LOCUS
DEFINITION Sequence 1009 from Patent WO0188188.
ACCESSION AX306258
VERSION AX306258.1 GI:17645513
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (sites)
Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
METHOD for examining ischemic conditions

JOURNAL Patent: WO 0188188-A 1009 22-NOV-2001;
FEATURES School Juridical Person Nihon University (JP)
  Location/Qualifiers
    1..483
      /organism="Mus musculus"
      /db_xref="taxon:10090"
BASE COUNT 127 a 139 c 121 g 96 t
ORIGIN

Query Match      52.4%; Score 417.8; DB 6; Length 483;
Best Local Similarity 93.2%; Pred. NO. 3.9e-120;
Matches 437; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 165 gtaaaaaagacattgagtgtatccaccctgggggcaatagccatattgccaatgg 224
DB 15 GTTAAAAAGACATTTGAGTGACGTAATCCACCCAGGGGCAATAGCCATATTGCCAATGG 74

QY 225 tgcggccgggtgtgtggcaacattacttcattgatgcaccctggaacccctgcggaagtgt 284
DB 75 TGCAGCCGGATGTGTGGCGACATTACTTTCATGATGCAGCCATGAATCCAGCGGAAGTCGT 134

QY 285 caagcagagagatgcagatgtacaaactcaccataccaccgggtgacagactgtgtacggc 344
DB 135 CAAGCAGAGGATGCGAGATGTACAACTCGCGGTACCACCGCGTGACAGACTGTGTTCGGGC 194

QY 345 agtgtggcaaatgaaggcgggcgccctttaccgcagctacacccacagctgacct 404
DB 195 AGTGTGGCAAAATGAAGGGCGCGGGCCCTTTTACCGCAGCTACACGACCAGCTGACCAT 254

QY 405 gaacgttcttcccaagccattcacttcattgatgcattccttcctgcaggagcatttaa 464
DB 255 GAATGTCCCTTCCAAAGCCATTACTTTCATGACCTATGAGTTCTCTGCAAGAGCAGCTTTAA 314

QY 465 ccccagagagcgtacacacccaagctcccaagctctctctgtgagcttgcagaggagctgt 524
DB 315 CCCCAGAGAGCGGTACAAACCCAGAGCTCCCATGTGCTGTGGGGCTCGCAGGAGCAGT 374

QY 525 agctcccgacgacacaccccaagctggagcttggaaaacactgctcaacacccagagctc 584
DB 375 AGCTCCCGCGCGCCACACCCCACTGAGAGCTTTGTTAAACACTGCTCAACACCCAGGAATC 434

QY 585 ctgtgcttgaactcacacattacagggacatatacagggcattggcgtagt 633
DB 435 CTGTGCTTTGAACCTCAAACTACTTACTGGACATCACAGGCATGGGTAGT 483

RESULT 12
AX071440 AX071440 401 bp DNA linear PAT 25-JAN-2001
LOCUS
DEFINITION Sequence 1912 from Patent WO0102568.
ACCESSION AX071440
VERSION AX071440.1 GI:12581791
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 401)
  Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
  Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
  Lanson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
  Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
  Human genes and gene expression products
  Patent: WO 0102568-A 1912 11-JAN-2001;
  CHIRON CORPORATION (US); HYSEQ, INC. (US)
FEATURES
  Location/Qualifiers
    1..401
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
  misc_feature 1..401
    /note="n = A,T,C or G"
BASE COUNT 100 a 104 c 120 g 76 t 1 others
```


ORIGIN

Query Match 49.1%; Score 392; DB 6; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.2e-112; Indels 0; Gaps 0;
Matches 392; Conservative 0; Mismatches 0;

Qy 8 gtctacagcctgacccagctgcccgcctatcgcaatgtgttgagggccctctggagatta 67
Db 9 gftctacagcctgacccagctgcccgcctatcgcaatgtgttgagggccctctggagatta 68
Qy 68 taagAACGAGGCGCTATGAGGCGCATGAGGGGCTGAACGCTCACAGCAACAGCGCGAG 127
Db 69 TAAGAACGAGGCGCTATGAGGCGCATGAGGGGCTGAACGCTCACAGCAACAGCGCGAG 128
Qy 128 ggcctgcccacgccccttatttgcctctactcgaagaaagttaaaaaagacattgagt 187
Db 129 GGCCTGCCACGCCCTTTATTTGGCTCTACTGAAAGATTAAAAAGACATTGAGTGATG 188
Qy 188 taatccacccctggggcgaatagccatattgcaatgtgtgcccgggtgtgtgccaacat 247
Db 189 TAATCCACCCCTGGGGCAATAGCCATATTGCCAATGGTGGCGGGGTGTGTGGCAACAT 248
Qy 248 tactctatgctgacgcccataaacctcgcgaagtgtgtcaagcagaggtgcatgtgtaca 307
Db 249 TACTTTCATGATGACCCATGAACCTGCGGAAGTGTCAAGCAGAGGATGCGAGATGTACA 308
Qy 308 actcaccataccacgggtgacagctgtgacggcagctgtgcaaaataaaggggcgg 367
Db 309 ACTCACCATTACACCGGTGACAGACTGTGTACGGGCGAGTGTGGCAAAATGAAGGGCGG 368
Qy 368 gggccttttacgcagctacacacccagctg 399
Db 369 GGGCCTTTTACGCGAGCTACACCACCCAGCTG 400

RESULT 13

AF288621 4112 bp mRNA linear ROD 27-OCT-2001
LOCUS Mus musculus mitochondrial solute carrier-like protein mRNA,
complete cds.

ACCESSION AF288621
VERSION AF288621.1 GI:16506177

KEYWORDS house mouse.
SOURCE Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Li,Q.-Z., Ruan,Q.-G., Eckenrode,S., Shi,J.-D., Cruz,P., Wang,C.-Y.
and She,J.-X.

TITLE A new gene which is highly expressed in NOD mice spleen

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 4112)

AUTHORS Li,Q.-Z., Shi,J.-D., Cruz,P., Echenrode,S., Ruan,Q.-G. and Wang,C.-Y.

TITLE Direct Submission

JOURNAL Submitted (20-JUL-2000) Pathology, University of Florida, 1600 SW

Archer Road, Rm. D6-15, Gainesville, FL 32610, USA

FEATURES

source Location/Qualifiers
1. 4112
/organism="Mus musculus"
/db_xref="taxon:10090"
116..1132
/note="50E12"
/codon_start=1
/product="mitochondrial solute carrier-like protein"
/protein_id="AA123859.1"
/db_xref="GI:16506178"

translation="MELRRGVGNQAGRWMDGDRDGGCKSDAGSEYENLPISAS
VSTHMTACAMAGILEHSIMTPVDSVKTRMOSLPDPAKRTSYIGALKRIMHTEGWRR
PLRLGNVMMGAGPAHAMYFACYENMKRTLNDVFSHGNSHLANGSWMATLLHDAY
MNPFAEYVKQLQYNSHQSAFSCIIRTWRTEGLGAFYRSYTTQLTMNIPFQSIHFIT

YEFLOEQVNPDRDYNPOSHIISGLAGALAAATAATPLDVCKTLINTONMALSIAVNS
GRISGMANAFRTVYQLNGLAGYFKGIQARVIYQMPSTAISWSVEFFKYLITRKOLEN
RTLY"

BASE COUNT 970 a 1024 c 970 g 1148 t
ORIGIN

Query Match 48.6%; Score 387.8; DB 10; Length 4112;
Best Local Similarity 69.3%; Pred. No. 1.3e-110;
Matches 543; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

Qy 1 atgcagagctctacagctgaccagctgcccgcctatcgcaatgtgttgagggccctctgg 60
Db 332 ATGCAGAGCTTTGAATCCAGATCCCAAGCCCGGTATACAAGCATCTATGCGCCCTCAAG 391
Qy 61 aggtattataaagaacggggcctctatgagggccctatgagggcctggaacgtcacagcaaca 120
Db 392 AGGATCATGCATGAAGGCTTCTGGAGGCCCTTGGCGGCCCTGAAGGTGATGATGATG 451
Qy 121 ggcgagggcctgcccacgccccttatttgcctgctacgaaagttaaaaaagacattg 180
Db 452 GGTGAGGGGCCCGGCACGCCCATGTATTTTGCCTGCTATGAAAACATGAAAAGACTTTA 511
Qy 181 agtgcgttaatccaccctggggcgaatagccatattgccaatgtgtgcggcggggtgtgtg 240
Db 512 AATGACGTTTTTACGCCACCAAGGAAACAGCCATCTAGCTAATGGGTAGCTGGGAGTATG 571
Qy 241 gcaacattactcatgatgacgccatgaacccctgcgaagtgtgcaagcagagagatgcag 300
Db 572 GCCACCCCTACTCCACGATGTCAGTAAATGAATCCAGCAGAAGTGGTGAACACACGGCTTAC 631
Qy 301 atgtacaactaccataccaccgggtgacagactgtgacgggcagtggtgcaaaatgaa 360
Db 632 ATGTACAACTCCAGACCCAGTACGCTTACGTTGTATCCGCGACAGTGTGGCGGACCGAG 691
Qy 361 ggggcccgggccccttttacgcgagctacacccagctgaccatgaacgttccttccaa 420
Db 692 GGGTTGGGGGCTTCTACAGGAGTTACACCACACAGCTGACCATGAATATCCCTTCCAG 751
Qy 421 gcaattcaattcatgacctatgaattcctgcaggagacacttaacccccagagacgttac 480
Db 752 TCAATTCACTTCATCAGCTATGAGTTTTCGAGGAGCAAGTCAACCCCTCGCCGGGACTAC 811
Qy 481 aacccaagctccacgtcctctctgtgagctgcgacgagctgtagctgcccagcagccaca 540
Db 812 AACCCACAGTCTCACATCATCTCAGGAGGCTTGGCCGGGACACTGGCCGACGCTGCCACC 871
Qy 541 acccactggagcttttgcaaaacactgctcaacccccagggagctgctgttggttgaac--- 597
Db 872 ACCCGCTGGAGCGTCTGCAAAACCCCTCTCAACACGCGAGGAGAACATGGCTCTCTCCCTG 931
Qy 598 tcacacattacaggaacatatcacaggcatggctagctgaccttcaggagcgggtatataagta 657
Db 932 GCCAACCTACAGCGCCGCGCTGTCGGGCATGGCCAAATGCTTCCGGACGGTGTACAGCTC 991
Qy 658 ggtggggtgacgcctatttccgaggggtgcaggcagagtaatttaccagatccccctcc 717
Db 992 AACGGCTTGGCGGCTATTTCAAAGGCATCCAGGCTCGAGTCATCTACCAATGCCCTTCC 1051
Qy 718 acagccatcgcatggtctgtgtatgagttcttcaaatcacctaatcactaaaaagcaagaa 777
Db 1052 ACCGCCATCTCCTGGTCGTTTATGAGTTCTTCAAGTACATCTTCAAGTACATCTTCAAGAGGAGCGAGCTG 1111
Qy 778 gag 780
Db 1112 GAG 1114

RESULT 14

AF223466
LOCUS Homo sapiens HT015 protein (HT015) mRNA, complete cds.
DEFINITION
ACCESSION AF223466
PRI 17-APR-2000

VERSION	AF223466.1	GI:7578782	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1429)		
AUTHORS	Song, H., Gao, G., Peng, Y., Ren, S., Chen, Z. and Han, Z.		
TITLE	A novel gene expressed in human hypothalamus		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1429)		
AUTHORS	Song, H., Gao, G., Peng, Y., Ren, S., Chen, Z. and Han, Z.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-JAN-2000) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, China		
FEATURES	Location/Qualifiers		
source	1..1429		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
gene	1..1429		
	/tissue_type="hypothalamus"		
CDS	45..1088		
	/gene="Ht015"		
	/codon_start=1		
	/product="Ht015 protein"		
	/protein_id="AAF64141.1"		
	/db_xref="GI:7578783"		
	/translation="MELRSGVGSQAVARRMGGDSRDGSGGKDATGSEDYENLPTSAS VSTMAGAGILEHSIMYPVDSVKTRMQLNDVPHOGNSHLANGSAGMATLLHDAY MNPAAVVKQLQYNSQHSRAISCIPTWRTGEGAFYRSTTQLTWNIPQSIHETI YEFLEQVNPHTYNPQSHIISGLAGALAAAATPLDVCKTLINTQENVALSANT GRLSGMANFRTVQLNGLPATSKASRRVSSTRCPPLGLSMSSSSTSPSASWKI ELHTKGRDHIFS"		
BASE COUNT	326 a 426 c 402 g 275 t		
ORIGIN			
Query Match	45.8%;	Score 365.6;	DB 9; Length 1429;
Best Local Similarity	68.7%;	Pred. No. 1.2e-103;	
Matches	533;	Conservative 0;	Mismatches 239; Indels 4; Gaps 2;
Qy	1	atgcagagtctacagcctgaccagctgcccgcctatgctgcaaatgtgttgagggccctctgg	60
Db	261	ATGCAGAGTTTGAGTCCAGATCCCAAGCCAGTACACAACTATCTACGGAGCCCTCAAG	320
Qy	61	aggattataagaacgaggggcctatgagggccatgaggggctgaaacgtcacagcaaca	120
Db	321	AAAATCATCGGACCGAAGGCTTCTGGAGGCCCTTGGAGGCGCTCAAGTCATGATCATG	380
Qy	121	ggcgaggcctccacgccccttatttgcctgctacagaaagttaaaaaacattg	180
Db	381	GGTCAGGCGCGCCGACATGCTATTTTCCCTGCTATGAAACATGAAAGGACTTTA	440
Qy	181	agtgatgtaattccaccctgggggcaatagccatatattgccaatggtggcgccgtgtgtg	240
Db	441	AATGACGTTTTCCACCAAGAAACAGACCCACCTAGCCACGGGATAGCTGGGAGTATG	500
Qy	241	gcacattactctatgatgagccatgaacccctcggaagtgggtcaagcagagatgcag	300
Db	501	GCCACCTCTCCACGATGCGGTATGAATCCAGCAGAAGTGGTGAAGCAGCGCTTGCG	560
Qy	301	atgtacaactcaccataccaccgggtacagactgtgtacgggcaggtggcacaatgaa	360
Db	561	ATGTACAACTCGACACCGGCTAGCAATCACTGTCATCCGACGGTGTGGAGGACCGAG	620
Qy	361	ggggccggggcctttaccgcagctacaccaccctgaccatgaacttcctttccaa	420
Db	621	GGGTGGGGGCTTCTACCGGAGGTACACACGAGCTGACCATGACATCCCTTCCAG	680
Qy	421	gccattcacttcagctatgaattcctgcaggagcactttaaccctccacagacggtac	480
Query Match	35.2%;	Score 281;	DB 10; Length 5869;
Best Local Similarity	70.1%;	Pred. No. 5.7e-77;	

```
Matches 393; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
Qy 223 ggtgcgcgggtgtgtggaacatttacttcattgatgcagccatgaacctctgcggaagtg 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2311 GGGTATAGCTGGGAGTATGGCCACCTACTCCACGATCAGTAATCCAGCAGCAAGTG 2370
Qy 283 gtcaagcagaggatgcagatgtacaaactcaccataccaccgggtgcagactgtgtacgg 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2371 GTGAACACAGCGCTTACAGATGTACAATCCAGCACCAGTCAGCCTTCAGTTGTATCCGG 2430
Qy 343 gcagtgtggcaaaatgaaggcgccgggcccctttaccgcagctacaccacccagctgacc 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2431 ACAGTGTGGCGGACCGAGGGGTGGGGCCCTTCTACAGGAGTTACACACACAGCTGACC 2490
Qy 403 atgaacgttctcttccaaagccattcaacttcattgatgacctatgaattcctgcaggagcacttt 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2491 ATGAATATCCCTCTTCCAGTCAATTCACCTTCACTATCATCCTATGAGTTTCTGCAGGAGCAAGTC 2550
Qy 463 aacccccagagacggttacaaacccaagctcccaagctcctctctgtgagcttgcgcaggagct 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2551 AACCTCGCCGGGACTACAAACCCACAGCTCTCACATCATCTCAGGAGGCCCTGGCGGGGCA 2610
Qy 523 gtagctgccgcagcccaacacccactggagcttgcacaaacactgctcaacacccagggag 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2611 CTGCCCGCAGCTGCCACACGCCCGCTGGAGCTCTGCAAAACCCCTCCTCAACACGCAAGGAG 2670
Qy 583 tccttgctttgaac---tcacacattacaggacatatcacaggcatggttagtgcttc 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2671 AACATGGCTCTCTCCCTGGCCCAAGTCAGCGCGCGGCTGTGCGGCATGGGCCAATGCGCTTC 2730
Qy 640 aggacggtatatcaagtaggtggtggtgaccgcctattccgaggggtgcagggccagagta 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2731 CGGACGGGTGTACCAGCTCAACGGCCTTGGCGGCTATTTCAAAGGCATCCAGGCTCGAGTC 2790
Qy 700 atttaccagatccccctccacagccatcgatggtctgtgtatgagttcttcaaatacctta 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2791 ATCTACCAGATGCCTTCCACCGCATCTCCTGGTCCGCTTTATGAGTTCTTCAAGTACATC 2850
Qy 760 atcaactaaaaggcaagaagag 780
    ||| ||| ||| ||| |||
Db 2851 CTTACAAAAGAGGACAGCTGGAG 2871
```

Search completed: September 28, 2002, 07:57:11
Job time: 14752 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 08:05:59 ; Search time 664.31 Seconds
(without alignments)
2062.437 Million cell updates/sec

Title: US-09-870-113-7
Perfect score: 798
Sequence: 1 atcgagagtctacagcctga.....agtggaggctggcaagtga 798

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	798	100.0	1294	22 AAI60661	Human polynucleoti
2	798	100.0	1322	22 AAF27733	Human transport pr
3	798	100.0	1336	21 AAC90457	Human uncoupling p
4	790	99.0	1316	22 AAI58875	Human polynucleoti
5	695	87.1	1418	21 AAC74843	Human OREF ORF398
6	639	80.1	710	20 AA215876	Human gene express
7	576	72.2	2502	22 AAS03906	Human secreted pro
8	556	69.7	2037	22 AAS62787	cDNA sequence #574
9	520	65.2	997	22 AAH22162	Human mitochondria

10	508	63.7	27960	22 AAK69779	Human immune/haema
11	508	63.7	27960	22 AAK73320	Human immune/haema
12	417.8	52.4	483	24 AAI99871	Mouse ischaemic co
13	392	49.1	401	22 AAF66156	Novel human polynu
14	376	47.1	1716	21 AAC77189	Human OREF ORF2744
15	365.6	45.8	1305	22 AAF59920	Human mitochondria
16	300	37.6	300	20 AA213026	Human gene express
17	281.2	35.2	1913	22 AAH89944	Human bone marrow
18	277.2	34.7	1814	22 AAK94871	Human full-length
19	276	34.6	1015	22 AAC92501	Human mitochondria
20	276	34.6	1673	22 AAK53156	Human polynucleoti
21	276	34.6	2562	22 AAH90057	Human bone marrow
22	276	34.6	2710	21 AAC66452	Human secreted pro
23	275.6	34.5	1890	22 AAK94897	Human full-length
24	272.2	34.1	600	22 AAH22163	Human mitochondria
25	267.6	33.5	1902	22 AAK52172	Human polynucleoti
26	250	31.3	8211	22 AAK80490	Human immune/haema
27	174	21.8	339	22 AAK58672	Human immune/haema
28	164.2	20.6	792	21 AAC77177	Human OREF ORF2732
29	161.4	20.2	1548	21 AAS65357	DNA encoding novel
30	159.6	20.0	504	23 AAS65354	DNA encoding novel
C 31	153.6	19.2	338	22 AAS23288	Human prostate can
32	141.4	17.7	2137	23 ABL08627	Drosophila melanog
33	138.2	17.3	975	21 AAC77173	Human OREF ORF2728
34	114.8	14.4	3028	22 AAC91320	Human polynucleoti
C 35	113.4	14.2	599	22 AAS24885	Human ovarian PCR-
C 36	113.4	14.2	599	22 AAH83525	Human ovarian tumo
37	105	13.2	842	22 AAK92357	Human cDNA 5'-end
38	105	13.2	842	22 AAK93895	Human cDNA clone r
39	105	13.2	2837	22 AAK80493	Human immune/haema
40	101.8	12.8	1468	21 AAC77175	Human OREF ORF2730
41	101	12.7	622	22 AAK91764	Human cDNA 5'-end
42	101	12.7	622	22 AAK93200	Human cDNA clone r
43	91.4	11.5	1758	20 AAV84573	Human secreted pro
44	91.4	11.5	1758	22 ABA83356	Human secreted pro
45	88	11.0	257	22 AAS31106	Human diagnostic a

ALIGNMENTS

RESULT 1
AAI60661/c
ID AAI60661 standard; cDNA; 1294 BP.
AC AAI60661;
XX
DT 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 4650.
XX
KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR P-PSDB: AAM41505.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX
XX Claim 1; SEQ ID NO 4650; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42113) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 1294 BP; 267 A; 335 C; 368 G; 324 T; 0 other;
SQ

Query Match 100.0%; Score 798; DB 22; Length 1294;
Best Local Similarity 100.0%; Pred. No. 2.5e-245;
Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctacagcctgagccagcctgcccgcctatgcgaatgtgttgagagccctctgg 60
DB 1134 ATGCAGAGCTACAGCCCTGAGCCAGCTGCCGCTATCCGATGTGTGGAGGCCCTCTGG 1075
QY 61 aggtattataagaagagggcctatggagccatgagggcgtgaacgtcacagcaaca 120
DB 1074 AGGATTATAAGAACGAGGGGCTATGGAGGCCATGAGGGGGCTGAACTCACAGCACACA 1015
QY 121 ggcgcagggcctgccacgcctcttatttgcctgctacgaaaagttaaaaaagacattg 180
DB 1014 GCGCGAGGGCCTGCCACGCCCTTTATTTTGCCTGCTACGAAAAAGTTAAAAAGACATTG 955
QY 181 agtgaatgaatccaccctgggggcaatagccatattgccaatgtgagccgggtgtgtg 240
DB 954 AGTGATGTAATCCACCCCTGGGGGCAATAGCCATATGTCGAATGCTGGCGCGGTGTGTG 895
QY 241 gcaactacttactatgatgcagccatgaacccctggaagtgtcaagcacagatgcag 300
DB 894 GCAACTTACTTATGATGACGACATGAACCCCTGGGAAGTGGTCAACGAGGATGCAG 835
QY 301 atgtacaactcaccataccacccgggtgacagactgtgtacgggagtggtgcaaaatgaa 360
DB 834 ATGTACAACTCACCATACCACCGGTGACAGACTGTGTACGGGAGTGTGCAAAATGAA 775
QY 361 ggggcccgggctttaccgcagctacacccaccagctgaccatgaacgttcccttccaa 420
DB 774 GGGGCGGGGGCTTTTACCAGAGCTACACCCAGCTGACCATGACGATTCCTTTCCAA 715
QY 421 gccattacttcagcctatgaattctctcagagacacttaacccccagagcgtac 480
DB 714 GCCATTCACTTCAGACCTATGAATTCCTGCAGGAGCACTTTAACCCCCAGAGCGGTAC 655
QY 481 aacccaagctcccacgtcctctctctgtgagcttgcgcagagagctgtagctgcccgcagccaca 540

Db 654 AACCCAGCTCCACGCTCTCTCTGGAGCTTGGCAGGAGCTAGCTGCCGAGCCACA 595
QY 541 acccaactggagctttgcaaaaaactgctcaaacccagggagctcttggtttgaactca 600
Db 594 ACCCCACTGGAGCTTGGCAAAACACTGCTCAACACCCAGGAGTCTTGGCTTTGAACCTCA 535
QY 601 cacattacagacatatcacagggcatggctagctccttcagagcgggtatatacaagtaggt 660
Db 534 CACATTACAGGACATATCACAGGCATGGCTAGTGGCTTCAGAGCGGTATATCAAGTAGGT 475
QY 661 ggggtgaccgcctatttccgaggggtgcagggccagagtaattaccagatccccctccaca 720
Db 474 GGGGTGACCGCTATTTCGAGGGGTGCAGGCCAGAGTAATTTACCAGATCCCCCTCCACA 415
QY 721 gccatcgcatgctgtgtatgatgagttcttcaaatcacataactaaaggaagcaagag 780
Db 414 GCCATCGCATGCTGTGTATGAGTCTTCAATACCTTAATCACTAAAGGCAAGAGAG 355
QY 781 tggagggcctgccaagtga 798
Db 354 TGGAGGGCTGGCAAGTGA 337

RESULT 2
AAF27733
ID AAF27733 standard; cDNA; 1322 BP.
XX
AC AAF27733;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33 coding sequence.
XX
KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200078953-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16668.
XX
PR 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI: 2001-041424/05.
DR P-PSDB: AAB60113.
XX
PT Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
XX Claim 5; Page 160; 165pp; English.
XX
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
SQ Sequence 1322 BP; 328 A; 376 C; 351 G; 267 T; 0 other;

```
Query Match      100.0%; Score 798; DB 22; Length 1322;
Best Local Similarity 100.0%; Pred. No. 2.6e-245;
Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctcacagcctgacccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
DB 186 atgcagagctcacagcctgacccagctgcccgcctatcgcaatgtgttgaggccctctgg 245

QY 61 aggattataaagaacgagggcctatggagcccatgaggggctgaacgtcacacaaca 120
DB 246 aggattataaagaacgagggcctatggagcccatgaggggctgaacgtcacacaaca 305

QY 121 ggcgagggcctgccacgccccttatttgcctgctacgaaagttaaaaaagacattg 180
DB 306 ggcgagggcctgccacgccccttatttgcctgctacgaaagttaaaaaagacattg 365

QY 181 agtgaatgaataccacccctgggggcaataagccatattgccaatgtgtcggcggtgtgtg 240
DB 366 agtgaatgaataccacccctgggggcaataagccatattgccaatgtgtcggcggtgtgtg 425

QY 241 gcaacattacttcattgatgcagccatgaacccctgcggaagtgttcaagcagaggatgcag 300
DB 426 gcaacattacttcattgatgcagccatgaacccctgcggaagtgttcaagcagaggatgcag 485

QY 301 atgtacaactaccataccacccgggtgcagactgtgtacgggaggtgtggcaaatgaa 360
DB 486 atgtacaactaccataccacccgggtgcagactgtgtacgggaggtgtggcaaatgaa 545

QY 361 ggggcccgggcccctttaccgagctcacaccacccagctgaccatgaacgttcttccaa 420
DB 546 ggggcccgggcccctttaccgagctcacaccacccagctgaccatgaacgttcttccaa 605

QY 421 gccattacttcattgatgcagccatgaacccctgcggaagtgttcaagcagaggatgcag 480
DB 606 gccattacttcattgatgcagccatgaacccctgcggaagtgttcaagcagaggatgcag 565

QY 481 aacccaagctccacgctcctctgtgagcttgcgagagctgtgctgagctgcgcagccaca 540
DB 666 aacccaagctccacgctcctctgtgagcttgcgagagctgtgctgagctgcgcagccaca 725

QY 541 acccactggagcgtttgcaaaacactgctcaacacccagagatccttggcttgaactca 600
DB 726 acccactggagcgtttgcaaaacactgctcaacacccagagatccttggcttgaactca 785

QY 601 cacattacaggacatatcacagggatggtgcttcagagcagcttatcaagttagt 660
DB 786 cacattacaggacatatcacagggatggtgcttcagagcagcttatcaagttagt 845

QY 661 ggggtgacgcgcctattccgaggggtgcagggcagagtaatttacagatccctccaca 720
DB 846 ggggtgacgcgcctattccgaggggtgcagggcagagtaatttacagatccctccaca 905

QY 721 gccatcgcatgctgtgtatagtttcttcaaatcccttaatacctaaaggcaagaagag 780
DB 906 gccatcgcatgctgtgtatagtttcttcaaatcccttaatacctaaaggcaagaagag 965

QY 781 tggagggctggcgaagtga 798
DB 966 tggagggctggcgaagtga 983
```

RESULT 3

AAC90457

ID AAC90457 standard; cDNA; 1336 BP.

XX AAC90457;

AC AAC90457;

XX 12-MAR-2001 (first entry)

DT Human uncoupling protein CDNA #6.

DE Human; uncoupling protein; immunosuppressive; antiarthritic;

XX Human; uncoupling protein; immunosuppressive; antiarthritic;

KW Human; uncoupling protein; immunosuppressive; antiarthritic;

antirheumatic; antiproliferative; cardiant; vasotropic;
cerebroprotective; neuroprotective, antibacterial; ophthalmological;
gastrointestinal; nephrotropic; gynaecological; vulnery; thrombolytic;
gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
infertility; ss.

XX Homo sapiens.

XX WO200061614-A2.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US09534.

XX 09-APR-1999; 99US-0128701.

XX 08-JUL-1999; 99US-0142821.

XX 18-AUG-1999; 99US-0149448.

XX 12-NOV-1999; 99US-0164751.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;

XX WPI: 2000-656322/63.

XX P-PSDB; AAB50383.

Uncoupling proteins and nucleic acid sequences encoding them, useful
for detecting, preventing and treating proliferative, neurological,
immune system, cardiovascular and gastrointestinal disorders -

XX Claim 1; Page 306-307; 343pp; English.

The present sequence is one of eighteen isolated nucleotide sequences
encoding uncoupling proteins. The nucleotide sequences may be used for
the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
or research purposes. The uncoupling protein encoded by the nucleotide
sequences may be used to treat disorders such as neural, immune,
muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
renal and proliferative disorders, wounds, infectious diseases,
thrombosis, arthritis, and infertility.

XX Sequence 1336 BP; 333 A; 377 C; 358 G; 268 T; 0 other;

Query Match 100.0%; Score 798; DB 21; Length 1336;
Best Local Similarity 100.0%; Pred. No. 2.6e-245;
Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctcacagcctgacccagctgcccgcctatcgcaatgtgttgaggccctctgg 60

DB 197 atgcagagctcacagcctgacccagctgcccgcctatcgcaatgtgttgaggccctctgg 256

QY 61 aggattataaagaacgagggcctatggagcccatgaggggctgaacgtcacacaaca 120

DB 257 aggattataaagaacgagggcctatggagcccatgaggggctgaacgtcacacaaca 316

QY 121 ggcgagggcctgccacgccccttatttgcctgctacgaaagttaaaaaagacattg 180

DB 317 ggcgagggcctgccacgccccttatttgcctgctacgaaagttaaaaaagacattg 376

QY 181 agtgaatgaataccacccctgggggcaatagccatattgccaatggtgcggcggtgtgtg 240

DB 377 agtgaatgaataccacccctgggggcaatagccatattgccaatggtgcggcggtgtgtg 436

QY 241 gcaacattacttcattgatgcagccatgaacccctgcggaagtgttcaagcagaggatgcag 300

DB 437 gcaacattacttcattgatgcagccatgaacccctgcggaagtgttcaagcagaggatgcag 496

QY 301 atgtacaactaccataccacccgggtgacagactgtgtacggcagctgtggcaaatgaa 360

DB 497 atgtacaactaccataccacccgggtgacagactgtgtacggcagctgtggcaaatgaa 556

Qy 361 ggggcccgggctttaccgcagctacaccacccagctgaccatgaacgttccttccaa 420
Db 557 ggggcccgggctttaccgcagctacaccacccagctgaccatgaacgttccttccaa 616
Qy 421 gccattcactcatgaactatgaattcctgcaggagacactttaaccccccagagcggtag 480
Db 617 gccattcactcatgaactatgaattcctgcaggagacactttaaccccccagagcggtag 676
Qy 481 aacccaagctccacgctctctggagcttgcgcagagctgtagctgcccagcgcaca 540
Db 677 aacccaagctccacgctctctggagcttgcgcagagctgtagctgcccagcgcaca 736
Qy 541 acccactggagctttgcaaacactgctcaacacccagagctcttggctttgaaactca 600
Db 737 acccactggagctttgcaaacactgctcaacacccagagctcttggctttgaaactca 796
Qy 601 cacattacaggacatatcacaggcatggtctagtgccctcaggagcgggtatatacagtaggt 660
Db 797 cacattacaggacatatcacaggcatggtctagtgccctcaggagcgggtatatacagtaggt 856
Qy 661 ggggtgaccgctatttccgaggggtgcaggccagagtaattaccagatccctccaca 720
Db 857 ggggtgaccgctatttccgaggggtgcaggccagagtaattaccagatccctccaca 916
Qy 721 gccatcgatggtctgtgtatgagttcttcaaatatcactaaaggcagaagag 780
Db 917 gccatcgatggtctgtgtatgagttcttcaaatatcactaaaggcagaagag 976
Qy 781 tggagggctggcaagtga 798
Db 977 tggagggctggcaagtga 994

RESULT 4

AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX AC AAI58875;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 1078.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX

DR WPI; 2001-442253/47.
DR P-PSDB; AAM39719.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 1078; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;

Query Match 99.0%; Score 790; DB 22; Length 1316;
Best Local Similarity 99.4%; Pred. No. 9.4e-243;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 atgcagagctcacagcctgaccagctgccgcctatgcgaatggttgaggccctctgg 60
Db 176 atgcagagctcacagctgaccagcctgcccagctggttcgaatggttgaggccctctgg 235
Qy 61 aggtattataagaacggaggccctatggaggcccatgagggggctgaaagtcacagcaaca 120
Db 236 aggtattataagaacggaggccctatggaggcccatgagggggctgaaagtcacagcaaca 295
Qy 121 ggcgcaggccctgccacgcctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 296 ggcgcaggccctgccacgcctttatttgcctgctacgaaagttaaaaaagacattg 355
Qy 181 agtcatgtaatccacctgggggcaatagccatattgccaatggtgcggccggtggtg 240
Db 356 agtcatgtaatccacctgggggcaatagccatattgccaatggtgcggccggtggtg 415
Qy 241 gcaacattacttcattgatgcagccatgaacccctgcgaagtgggtcaagcagagatgcag 300
Db 416 gcaacattacttcattgatgcagccatgaacccctgcgaagtgggtcaagcagagatgcag 475
Qy 301 atgtacaactacacataccacccgggtgacagactgtgtacgggagtggtggcaaatgaa 360
Db 476 atgtacaactacacataccacccgggtgacagactgtgtacgggagtggtggcaaatgaa 535
Qy 361 ggggcccgggcctttaccgcagctacaccacccagctgaccatgaacgttccttccaa 420
Db 536 ggggcccgggcctttaccgcagctacaccacccagctgaccatgaacgttccttccaa 595
Qy 421 gccattcacttcacgtatgaattcctgcaggagcactttaaccccccagagcggtag 480
Db 596 gccattcacttcacgtatgaattcctgcaggagcactttaaccccccagagcggtag 655
Qy 481 aacccaagctccacgctctctctctgcaggcttgcgcagagcgttagctgcgcagcaca 540
Db 656 aacccaagctccacgctctctctctgcaggcttgcgcagagcgttagctgcgcagcaca 715
Qy 541 acccactgagcgtttgcaaacactgctcaacacccagagctcttggctttgaaactca 600
Db 716 acccactgagcgtttgcaaacactgctcaacacccagagctcttggctttgaaactca 775
Qy 601 cacattacaggacatatcacaggcatggtcttccatcaggagcgggtatatacagtaggt 660

Db 776 cacattacagcacatcacagcctgctgtagtgccttcaggagcgtatatacaagtagt 835
 Qy 661 ggggtgaccgcctattccagagggtgcagccagagtaattaccagatccctccaca 720
 Db 836 ggggtgaccgcctattccagagggtgcagccagagtaattaccagatccctccaca 895
 Qy 721 gccatgcgatgctgtgatgattcttcaaatcacctaatcaactaaaggcagaagag 780
 Db 896 gccatgcgatgctgtgatgattcttcaaatcacctaatcaactaaaggcagaagag 955
 Qy 781 tggagggtgccaagtga 798
 Db 956 tggagggtgccaagtga 973

RESULT 5

AAC74843
 ID AAC74843 standard; cDNA; 1418 BP.
 XX
 AC AAC74843;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF398 polynucleotide sequence SEQ ID NO:795.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; vasotrophic; antidiabetic;
 KW immunostimulant; thrombolytic; coagulant; immunosuppressive; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antidiabetic;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 XX WPI; 2000-602362/57.
 DR P-PSDB; AAB40634.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame x,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 5; Page 853-854; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotrophic;

CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergic, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 1418 BP; 353 A; 400 C; 368 G; 297 T; 0 other;

Query Match 87.1%; Score 695; DB 21; Length 1418;
 Best Local Similarity 89.6%; Pred. No. 2.9e-212;
 Matches 798; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

Qy 1 atgcagagtctacagcctgacccagctgcgcgctatcgcaatgtgttgaggccctctg 60
 Db 178 atgcagagtctacagcctgacccagctgcgcgctatcgcaatgtgttgaggccctctg 237
 Qy 61 agattataaagcagaggccctatggagccatgaggggctgaacgtcacagcaaca 120
 Db 238 agattataaagcagaggccctatggagccatgaggggctgaacgtcacagcaaca 297
 Qy 121 ggcgaggcctgccacgcctttatttgcctgctacgaaagttaaaaagacattg 180
 Db 298 ggcgaggcctgccacgcctttatttgcctgctacgaaagttaaaaagacattg 357
 Qy 181 agtgatgataccacctgggggcaatggccatattgccaatggtgcggccggtgtgtg 240
 Db 358 agtgatgataccacctgggggcaatggccatattgccaatggtgcggccggtgtgtg 417
 Qy 241 gcaacattacttcatgctgacgcatgaacctgcgga----- 278
 Db 418 gcaacattacttcatgctgacgcatgaacctgcggaagcgtgctgctgactgggg 477
 Qy 279 ----- 278
 Db 478 ctctgaatctgatactctccatcacccggttggtgctgtcaccattctcttcctgtg 537
 Qy 279 -----agtgtcaagcagaggatgagatgtacaaactacacacacgggtg 327
 Db 538 atggcactactagtgtcaagcagaggatgagatgtacaaactacacacacgggtg 597
 Qy 328 acagactgttacgggcagtggtgcacaaatgaaggccgagggccttttacgcagctac 387
 Db 598 acagactgttacgggcagtggtgcacaaatgaaggccgagggccttttacgcagctac 657
 Qy 388 accaccagctgacatgaacgcttcttccaaagccattcacttcatgacctatgaattc 447
 Db 658 accaccagctgacatgaacgcttcttccaaagccattcacttcatgacctatgaattc 717
 Qy 448 ctgcaggagcatttaacccccagagacggtacacacccaagctccacgctctctctgga 507
 Db 718 ctgcaggagcatttaacccccagagacggtacacacccaagctccacgctctctctgga 777
 Qy 508 gcttgcgagagcgtgtagctgcgcagcagcaacacccacacggagctttgcaaacactg 567
 Db 778 gcttgcgagagcgtgtagctgcgcagcagcaacacccacacggagctttgcaaacactg 837
 Qy 568 ctcaacacccagagctctcttggtcttgtaactcacacattacaggacatatacacaggcatg 627
 Db 838 ctcaacacccagagctctcttggtcttgtaactcacacattacaggacatatacacaggcatg 897
 Qy 628 gctagtgcttcagagcaggtatatacaagtagtggtgggtgacgcgctatttccaggggtg 687

Db 898 gtagtgcttcaggacggtatatcaagtagtgggtgaccgcctatttccgggggtg 957
Qy 688 caggccagagtaattaccagatccctccacagccatcgcatgctgtgtatgagttc 747
Db 958 caggccagagtaattaccagatccctccacagccatcgcatgctgtgtatgagttc 1017
Qy 748 ttcaataacctaatcaactaaaggcaagaagtagtgagggtggcgaagtga 798
Db 1018 ttcaataacctaatcaactaaaggcaagaagtagtgagggtggcgaagtga 1068

RESULT 6
AAZ15876
ID AAZ15876 standard; cDNA; 710 BP.
XX
AC AAZ15876;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:3345.
XX
KW Human; gene: gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN W09938972-A2.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-US01619.
XX
PR 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE) HYSEQ INC.
XX
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lanson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX
DR WPI; 1999-494092/41.
XX
PT Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX
PS Claim 1; Page 1602; 2479pp; English.
XX

The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ1779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in AAZ12532 to AAZ1779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the

CC diagnosis, prognosis and management of colorectal cancer, breast cancer, CC and lung cancer. The polynucleotides can also be used to screen for CC peptide analogues and antagonists.
XX
SQ Sequence 710 BP; 186 A; 187 C; 183 G; 154 T; 0 other;

Query Match 80.1%; Score 639; DB 20; Length 710;
Best Local Similarity 100.0%; Pred. No. 1.9e-194;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 gaaaagttaaaaaagacattgagtgatgaatccacccctgggggcaatagccattgtgcc 219
Db 46 gaaaagttaaaaaagacattgagtgatgaatccacccctgggggcaatagccattgtgcc 105
Qy 220 aatggtgcggccgggtgtgtggcaacatttacttctatgatcagccatgaacccctgcgaa 279
Db 106 aatggtgcggccgggtgtgtggcaacatttacttctatgatcagccatgaacccctgcgaa 165
Qy 280 gtggtcaagcagagatgcagatgtacaactcacatcacccgggtgacagactgtgta 339
Db 166 gtggtcaagcagagatgcagatgtacaactcacatcacccgggtgacagactgtgta 225
Qy 340 cgggcagtggtggcaaaatgaaggccgggcccctttaccgcagctacacccaccagctg 399
Db 226 cgggcagtggtggcaaaatgaaggccgggcccctttaccgcagctacacccaccagctg 285
Qy 400 accatgaacgttcccttccaaagccattcacttgcctatgacctatgaattcctcagagcac 459
Db 286 accatgaacgttcccttccaaagccattcacttgcctatgacctatgaattcctcagagcac 345
Qy 460 tttacccccagagcgtgtacaaccccaagctccacagctccctctctgagcttgcgcagga 519
Db 346 tttacccccagagcgtgtacaaccccaagctccacagctccctctctgagcttgcgcagga 405
Qy 520 gctgtagctgcgcagccacacaccccccactggacgttttgcaaaacactgctcaacacccag 579
Db 406 gctgtagctgcgcagccacacaccccccactggacgttttgcaaaacactgctcaacacccag 465
Qy 580 gagtccctggcttgaactcaacattacagagacatatacacagcagctggtagtgcttc 639
Db 466 gagtccctggcttgaactcaacattacagagacatatacacagcagctggtagtgcttc 525
Qy 640 aggcaggtatatcaagtagtggtgggtgacccctatttccaggggtgcagggccagagta 699
Db 526 aggcaggtatatcaagtagtggtgggtgacccctatttccaggggtgcagggccagagta 585
Qy 700 attaccagatccccctccacagccatgcgcatggtgtgtatgagttcttcaataacctta 759
Db 586 attaccagatccccctccacagccatgcgcatggtgtgtatgagttcttcaataacctta 645
Qy 760 atcaactaaaaggcaagaagtagtgagggtggcgaagtga 798
Db 646 atcaactaaaaggcaagaagtagtgagggtggcgaagtga 684

RESULT 7
AAS03906
ID AAS03906 standard; cDNA; 2502 BP.
XX
AC AAS03906;
XX
DT 29-AUG-2001 (first entry)
XX
DE Human secreted protein gene #25.
XX
KW Human secreted protein; autoimmune disorder; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;
KW nervous system disorder; bacterial infection; viral infection; ss;
KW fungal infection; ocular disorder; wound healing; tissue regeneration;
KW epithelial cell proliferation; skin ageing; chemotaxis; I9G FC region.
XX
OS Homo sapiens.

XX WO200123598-A1.
PN 05-APR-2001.
XX 26-SEP-2000; 2000WO-US26324.
PF 27-SEP-1999; 9905-US155807.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Komatsoulis G, Ruben SM, Rosen CA;
PI WPI; 2001-281684/29.
DR P-PSDB; AAU01950, AAU01989.
XX Forty one nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
XX Disclosure: Page 460-461; 518pp; English.
XX Sequences AAS03873-AAS03922 represent isolated nucleic acid molecules and
CC PCR primers of the invention. acid of the invention. Secreted proteins
CC and their related nucleic acids can be used in the diagnosis of or
CC susceptibility to a pathological condition by determining the presence or
CC absence of a mutation in a nucleic acid or the presence or amount of
CC expression of a secreted protein. The sequences are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. The antibodies to the
CC polypeptides can also be used in alleviating symptoms associated with
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). The disorders include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The peptides can also be used to aid wound healing and epithelial cell
CC proliferation, to help prevent skin ageing due to sunburn, to maintain
CC organs before transplantation, to regenerate tissues, in chemotaxis and
CC as a food additive or preservative to alter storage capabilities.
XX
XX Sequence 2502 BP; 587 A; 655 C; 669 G; 589 T; 2 other;
SQ

Query Match 72.2%; Score 576; DB 22; Length 2502;
Best Local Similarity 100.0%; Pred. No. 5.7e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 ggtcgccgggtgtgtggaacattactctatgatgcagccatgaacccctgcggaagt 282
Db 1513 ggtcgccgggtgtgtggaacattactctatgatgcagccatgaacccctgcggaagt 1572

QY 283 gtcaagcagagatgcagatgtacaactcaccatcaccacccgggtgacagactgtgtacgg 342
Db 1573 gtcaagcagagatgcagatgtacaactcaccatcaccacccgggtgacagactgtgtacgg 1632

QY 343 gcagtggtgcaaaatgaagggccgggccccttttaccgcagctacaccaccagctgacc 402
Db 1633 gcagtggtgcaaaatgaagggccgggccccttttaccgcagctacaccaccagctgacc 1692

QY 403 atgaacgttcctttccaaagccattcaactctatgcacctatgaattcctgcagagcacctt 462
Db 1693 atgaacgttcctttccaaagccattcaactctatgcacctatgaattcctgcagagcacctt 1752

QY 463 aacccccagagcgtacaacccaagctcccaactcctctctgtgagcttgcgcaggact 522
Db 1753 aacccccagagcgtacaacccaagctcccaactcctctctgtgagcttgcgcaggact 1812

QY 523 gtagtgcgcagcagcacacccccactgacgttttgcaaaaactcactcacaacccagag 582
Db 1813 gtagtgcgcagcagcacacccccactgacgttttgcaaaaactcactcacaacccagag 1872

QY 583 tcttggtcttgaactcacacattacagacacatatacacagcatggtgtagtccttcagg 642
Db 1873 tcttggtcttgaactcacacattacagacacatatacacagcatggtgtagtccttcagg 1932

QY 643 acggtatatcaagtagtggtgacccgctatttccagggggtgcagccagagtaatt 702
Db 1933 acggtatatcaagtagtggtgacccgctatttccagggggtgcagccagagtaatt 1992

QY 703 taccagatccctccacagcaccatgcgtgtgtgtatgatttcttcaaatacctaattc 762
Db 1993 taccagatccctccacagcaccatgcgtgtgtgtatgatttcttcaaatacctaattc 2052

QY 763 actaaaaggcaagaagagtgaggggctggcgaagtga 798
Db 2053 actaaaaggcaagaagagtgaggggctggcgaagtga 2088

RESULT 8
AAS62787
ID AAS62787 standard; cDNA; 2037 BP.
XX
AC AAS62787;
XX
DT 14-FEB-2002 (first entry)
XX
DE cDNA sequence #574 encoding novel human secreted protein.
XX
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW immunosuppressive; antirheumatic; ss.
XX
OS Homo sapiens.
XX
PN WO200177291-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US10485.
XX
PR 06-APR-2000; 2000US-195604P.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulikota K, Graham JR;
XX
DR WPI; 2002-010900/01.
XX
PT New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease -
XX
PS Claim 1; Page 364-365; 391pp; English.
XX
CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides
CC a method for producing proteins from these polynucleotide sequences.
CC The proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are
CC useful for treating diseases such as hyperproliferative disorders
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory
CC disorders (e.g. arthritis) and infectious disorders (e.g. hepatitis).
CC The polynucleotide sequences of the invention are also useful in gene
CC therapy. AAS62214-AAS62838 represent the cDNA sequences of the
CC invention that encode for novel human secreted proteins.
XX
SQ Sequence 2037 BP; 475 A; 522 C; 524 G; 515 T; 1 other;

```
Query Match      69.7%; Score 556; DB 24; Length 2037;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 570; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 223 ggtgcgcgggtgtgtggaacattacttcatgatgcagccatgaaccctgcggaagt 282
Db 1138 ggtgcgcgggtgtgtggaacattacttcatgatgcagccatgaaccctgcggaagt 1197

Qy 283 gtcaagcagagatgcagatgtacaactcacatcacaccgggtgacagctgtgtacg 342
Db 1198 gtcaagcagagatgcagatgtacaactcacatcacaccgggtgacagctgtgtacg 1257

Qy 343 gcagtgtgcaaaatgaaggccggcggccttttacgcagctacacaccagctgacc 402
Db 1258 gcagtgtgcaaaatgaaggccggcggccttttacgcagctacacaccagctgacc 1317

Qy 403 atgaagcttcttccaagccattcaacttcatgacctatgaattcctgcagagcacttt 462
Db 1318 atgaagcttcttccaagccattcaacttcatgacctatgaattcctgcagagcacttt 1376

Qy 463 aacccccagagacggtacaaaccccaagctccacgtctctctgtgagcttgcgcggagct 522
Db 1377 aacccccagagacggtacaaaccccaagctccacgtctctctgtgagcttgcgcggagct 1436

Qy 523 gtactgtcgcagccacacacccccactggacgttttgcaaaacactgctcaacaccaggag 582
Db 1437 gtactgtcgcagccacacacccccactggacgttttgcaaaacactgctcaacaccaggag 1496

Qy 583 tcttgggttgaactcacattacagggacatatacagcagctggtgcttgccttcagg 642
Db 1497 tcttgggttgaactcacattacagggacatatacagcagctggtgcttgccttcagg 1556

Qy 643 acggtatataagtaggtggtggtgacgcctattttccgaggggtgcagccagagtaatt 702
Db 1557 acggtatataagtaggtggtggtgacgcctattttccgaggggtgcagccagagtaatt 1616

Qy 703 taccagatccccctccacagcctgcgtgctgtgtatgattcttcaaataccctaate 762
Db 1617 taccagatccccctccacagcctgcgtgctgtgtatgattcttcaaataccctaate 1676

Qy 763 actaaaggccaagaagtagtgagggtgccaagtga 798
Db 1677 actaaaggccaagaagtagtgagggtgccaagtga 1712

RESULT 9
AAH22162
ID AAH22162 standard; cDNA; 997 BP.
XX
AC AAH22162;
XX
XX 20-AUG-2001 (first entry)
DT
DE Human mitochondrial solute carrier (hMSC-homologue) encoding cDNA SEQ:6.
XX
XX Human; hMSC; hMSC-homologue; mitochondrial solute carrier; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 379..909
CDS /*tag= a
FT /product= "mitochondrial solute carrier (hMSC-homologue)"
FT
XX
XX CN1281895-A.
FN
XX
XX 31-JAN-2001.
PD
XX
XX 27-JUN-2000; 2000CN-0116795.
PF
XX
XX 27-JUN-2000; 2000CN-0116795.
PR
```

```
XX (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
PA Li N, Xu X, Xiao H;
XX
XX WPI; 2001-282781/30.
DR P-PSDB; AAB98221.
XX
XX New human mitochondrial solute carrier (hMSC) protein isomer and its
PT code sequence -
XX
XX Claim 1; Page 17-18 (disclosure); 22pp; Chinese.
XX
XX The present invention describes a human mitochondrial solute carrier
CC protein, designated hMSC-homologue, which is expressed in human
CC hypophysis. Also described are methods for the preparation and
CC detection of hMSC-homologue protein and nucleotide sequences. The
CC present sequence encodes hMSC-homologue, as given in the present
CC invention.
XX
XX Sequence 997 BP; 240 A; 261 C; 242 G; 254 T; 0 other;
```

```
Query Match      65.2%; Score 520; DB 22; Length 997;
Best Local Similarity 100.0%; Pred. No. 3.3e-156;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 279 agtggtcaagcagagatgcagatgtacaactcacatcacaccgggtgacagactgtgt 338
Db 390 agtggtcaagcagagatgcagatgtacaactcacatcacaccgggtgacagactgtgt 449

Qy 339 acgggcagtggtgcaaaatgaaggccggcggccttttacgcagctacacaccagct 398
Db 450 acgggcagtggtgcaaaatgaaggccggcggccttttacgcagctacacaccagct 509

Qy 399 gaccatgaagcttcttccaagccattcacttcatgacctatgaattcctgcaggagca 458
Db 510 gaccatgaagcttcttccaagccattcacttcatgacctatgaattcctgcaggagca 569

Qy 459 ctttaacccccagagacggtacaaaccccaagctccacgtctctctgtgagcttgcgagg 518
Db 570 ctttaacccccagagacggtacaaaccccaagctccacgtctctctgtgagcttgcgagg 629

Qy 519 agctgtagctgcgcagccacacaccccaactggacgttttgcaaaacactgctcaacacca 578
Db 630 agctgtagctgcgcagccacacaccccaactggacgttttgcaaaacactgctcaacacca 689

Qy 579 ggagtccttggcttgaactcacacattacaggacatatacagggcatggctagtgcctt 638
Db 690 ggagtccttggcttgaactcacacattacaggacatatacagggcatggctagtgcctt 749

Qy 639 caggacggtatatacagtagtggtgggtgacgcctattttccgaggggtgcaggccagagt 698
Db 750 caggacggtatatacagtagtggtgggtgacgcctattttccgaggggtgcaggccagagt 809

Qy 699 aatttaccagatccccctccacagcctgcagctggtgtgtgtatgattcttcaataacct 758
Db 810 aatttaccagatccccctccacagcctgcagctggtgtgtgtatgattcttcaataacct 869

Qy 759 aatcactaaaggccaagaagtagtgagggtgccaagtga 798
Db 870 aatcactaaaggccaagaagtagtgagggtgccaagtga 909
```

```
RESULT 10
AAK69779
ID AAK69779 standard; DNA; 27960 BP.
XX
AC AAK69779;
XX
XX 06-NOV-2001 (first entry)
DT
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24591.
DE
```

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX DR WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
XX Disclosure; SEQ ID NO 24591; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;

Query Match 63.7%; Score 508; DB 22; Length 27960;
Best Local Similarity 99.8%; Pred. No. 1.3e-151;
Matches 519; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Oy 279 agtggtaagcagagatgcagatgtacaaactcaccatcaccacccgggtgacagactgtgt 338
Db 20845 agtggtaagcagagatgcagatgtacaaactcaccatcaccacccgggtgacagactgtgt 20904
Oy 339 acgggcagtggtgcaaaatgaaggccggcgcccttttacccgagctacacacccagct 398
Db 20905 acgggcagtggtgcaaaatgaaggccggcgcccttttacccgagctacacacccagct 20963
Oy 399 gaccatgaacgtctcttccaaagccattcacttcacatgcattgaattcctgcaggagca 458
Db 20964 gaccatgaacgtctcttccaaagccattcacttcacatgcattgaattcctgcaggagca 21023
Oy 459 ctttaaccccgagagcgttacaaaccccaagctcccaagctctctctgagcttgcgagg 518
Db 21024 ctttaaccccgagagcgttacaaaccccaagctcccaagctctctctgagcttgcgagg 21083
Oy 519 agctgtagctccgagcagcaacaccccaagctgagcttgcagaaactgctcaacacca 578
Db 21084 agctgtagctccgagcagcaacaccccaagctgagcttgcagaaactgctcaacacca 21143
Oy 579 ggaagctcttgcttgaactcacacattacagggacatcatcagggcagtggttagtgcctt 638
Db 21144 ggaagctcttgcttgaactcacacattacagggacatcatcagggcagtggttagtgcctt 21203
Oy 639 caggacggtatatcaagtagtggtgagccgctatttcagaggggtgcaggccagagt 698
Db 21204 caggacggtatatcaagtagtggtgagccgctatttcagaggggtgcaggccagagt 21263
Oy 699 aattaccagatccctccacagccatcgatgctgtgtatagttcttcaataacct 758
Db 21264 aattaccagatccctccacagccatcgatgctgtgtatagttcttcaataacct 21323
Oy 759 aatcactaaaaggcaagaagatggagggtggaagtgga 798

Db 21324 aatcactaaaaggcaagaagatggagggtggaagtgga 21363
|||||
RESULT 11
AAK73320
ID AAK73320 standard; DNA; 27960 BP.
XX
XX AC AAK73320;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28132.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209457.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
DR
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 28132; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients' own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;

Query Match 63.7%; Score 508; DB 22; Length 27960;
Best Local Similarity 99.8%; Pred. No. 1.3e-151;
Matches 519; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 279 agtgggtcaagcaggatgcagatgtacaactcacaccacccgggtgacagactgtgt 338
|||||
Db 20845 agtgggtcaagcaggatgcagatgtacaactcacaccacccgggtgacagactgtgt 20904
QY 339 acggggcagtgtggcaaaatgaaggggcgccggttttacgcagctacacacccagct 398
|||||
Db 20905 acggggcagtgtggcaaaatgaaggggcgccggttttacgcagctacacacccagct 20963
QY 399 gacctatgaacgttcctttccaagccattcaactcatgatgcattaatcttcgcaggagca 458
|||||
Db 20964 gacctatgaacgttcctttccaagccattcaactcatgatgcattaatcttcgcaggagca 21023
QY 459 cttaaaccgccagagacggtgataaacacccaggctcccacgtctctcttgagcttgcgcagg 518
|||||
Db 21024 cttaaaccgccagagacggtgataaacacccaggctcccacgtctctcttgagcttgcgcagg 21083
QY 519 agctgtagctgcgcagaggccacacccccactggacgttttgaaaaactgtctcaacacca 578
|||||
Db 21084 agctgtagctgcgcagaggccacacccccactggacgttttgaaaaactgtctcaacacca 21143

PS Claim 9; Page 818; 1046pp; English.

XX The present sequence is one of 3351 sequences in a library of human
XX polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.

XX Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match 49.1%; Score 392; DB 22; Length 401;

Best Local Similarity 100.0%; Pred. No. 2.3e-115;

Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 gtctacagcctgacccagctgcccgtatcgcaatgtgttgaggccctctggaggatta 67

Dy 9 gtctacagcctgacccagctgcccgtatcgcaatgtgttgaggccctctggaggatta 68

Qy 68 taagaacaggaggcctatggaggcccatgaggggcctgaacgtcacacacagcgcgag 127

Dy 69 taagaacaggaggcctatggaggcccatgaggggcctgaacgtcacacacagcgcgag 128

Qy 128 ggcctgccacgcctttatttgcctctacgaaagttaaaagacattgagtgatg 187

Dy 129 ggcctgccacgcctttatttgcctctacgaaagttaaaagacattgagtgatg 188

Qy 188 taatcacccctgggggcaatagccatattgccaattggtgcggccgggtgtgtggcaacat 247

Dy 189 taatcacccctgggggcaatagccatattgccaattggtgcggccgggtgtgtggcaacat 248

Qy 248 tacttcgatgacagcctgaacctgcggaagtgttcaacagagagtgagatgtaca 307

Dy 249 tacttcgatgacagcctgaacctgcggaagtgttcaacagagagtgagatgtaca 308

Qy 308 actcaccataccacgggtgacagactgtacggcgagtggtgcgaataaagggcgcg 367

Dy 309 actcaccataccacgggtgacagactgtacggcgagtggtgcgaataaagggcgcg 368

Qy 368 gggccttttacgcagctacaccccccagctg 399

Dy 369 gggccttttacgcagctacaccccccagctg 400

RESULT 14

AAC77189

ID AAC77189 standard; cDNA; 1716 BP.

XX AAC77189;

AC AAC77189;

XX 08-FEB-2001 (first entry)

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF2744 polynucleotide sequence SEQ ID NO:5487.

DE Human;

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;

XX vulnerable; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;

XX anticonvulsant; osteopathic; antithrombotic; immunosuppressant; cardiac;

XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

XX hypotensive; dermatological; immunosuppressive; antiinflammatory;

XX antitumor; antibacterial; antifungal; antineoplastic; antithyroid;

XX antineoplastic; gene therapy; cancer; proliferative disorder; hypertension;

XX neurodegenerative disorder; osteoarthritis; graft vs host disease;

XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive; ss.

OS Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB42980.

XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX Claim 5; Page 4661-4662; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antiproliferative; antiparkinsonian; neurotropic; immunoprotective;
XX osteopathic; anticonvulsant; antithrombotic; immunosuppressant;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; antineoplastic;
XX antitumor; antibacterial; antifungal; antineoplastic; antithyroid;
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus, lupus
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 1716 BP; 411 A; 457 C; 446 G; 402 T; 0 other;

Query Match 47.1%; Score 376; DB 21; Length 1716;

Best Local Similarity 68.7%; Pred. No. 6.7e-110;

Matches 533; Conservative 0; Mismatches 240; Indels 3; Gaps 1;

Qy 1 atgcagagctctacagcctgacccagcgtcccgctctatcgcaatgtgttgaggccctctgg 60

Dy 256 atgcagagcttgcctccagcctcccaagcccgatcacagatctatcagcgccctcaag 315

Qy 61 aggattataaagcagcggcctatggaggccatgagggcgtgaacgtcacacacaca 120

Dy 316 aaaaatcgagcagcggcctctgaggccctctgaggcgtgaacgtcacacacacag 375

Qy 121 ggcgcagcggcctgcccagcctttatttgcctgctacgaaagttaaaagacattg 180

Dy 376 ggtgcagggccagcccgatgcatgtatttgcctgctacgaaagttaaaagacattg 435

181	Qy	agtgatgtaattccacacctgggggcaatagccatttgccaatggtgcggccgggtgtgtg	240
436	Db	aatgagctttccacaccagaagaaacgcacctgaccaacgggatagctgggagatgtg	495
241	Qy	gcaacattacttcattgatgcagccatgaacctgcggaagtgtcaagcagagagatgcag	300
496	Db	gcacacctgtccacatgcggttaatgaattccagcagaagtgtgaagcagcgctgcgag	555
301	Qy	atgtacaactcacatcacaccgggtgacagactgtgtacggcgacgtgtggcacaatgaa	360
556	Db	atgtacaactcgacgacccggtcagcaatcagctgcataccggacggtgtggaggaccgag	615
361	Qy	ggggcgccggccttttacccgagctacacacccacgctgacctgaaccttcctttccaa	420
616	Db	gggtggggcctttaccggagctacaccacgcagctgacctgaacctccctcccgag	675
421	Qy	gccattcaactcatgacctatgaattcctcgcaggagacactttaacccccagagacggtac	480
676	Db	tccatccactctcatcctatggttcctcgcaggagcagtgtaacccccaccacggacctac	735
481	Qy	aaccacagctcccaacgtccctctcttgagcttgcgcaggagctgtagctgccgcagccaa	540
736	Db	aaccgcgagtcaccaatcatctcagcgggctgcgcggccctgcgcggccgcgcagc	795
541	Qy	accacctgacgtttgtgcaaaacactgctcaacacccagagatccttggc---tttgaac	597
796	Db	accacctggacgtctgtgaagacccctctgaacactcaggagagcgtggccctctcgctg	855
598	Qy	tcacacattacaggacatatcacaggcatggtctagcttccttcaggacggtatatcaagta	657
856	Db	gccaacatcagcgccgcgtgtcgggtatggccaatgccttcgcgcaggtgtaccagctc	915
658	Qy	ggctgggtgacgcacctattcccaggggtgcaggccagagaaattaccagatccccctcc	717
916	Db	aacggcctggccggctacttcaaaaggcatccagcgcggtgtcatctaccagatgcacctcc	975
718	Qy	acagccaatcgcatggctgtgtgtatgagttcttcaaatcactaatcactaaaaggca	773
976	Db	acagcacttcttgctgtctctatgagttcttcaaatcttctccacgaagccca	1031

RESULT	15	
AAF59920		
ID	AAF59920	standard; cDNA; 1305 BP.
XX	AC	
XX	AAAF59920;	
XX	04-MAY-2001	(first entry)
XX	Human mitochondrial	solute carrier protein hMSC-o cDNA..
XX	Human mitochondrial	solute carrier protein; hMSC-o; hypothalamus;
KW	preparation; detection;	ss.
XX	Homo sapiens.	
OS	CN1269409-A.	
XX	11-OCT-2000.	
PD	17-MAR-2000;	2000CN-0114958.
XX	17-MAR-2000;	2000CN-0114958.
PF	(SREN-)	SOUTHERN RES CENT NAT HUMAN GENE GROUP.
XX	Zhang X,	Gao X, Xiao H;
XX	WPI; 2001-050544/07.	
XX	P-PSDB; AAB60658.	
XX	New human mitochondrial	solute carrier protein and its nucleic acid
PT		
XX		

Search completed: September 28, 2002, 08:06:20
Job time: 11575 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:55:43 ; Search time 182.06 Seconds
(without alignments)
1076.653 Million cell updates/sec

Title: US-09-870-113-7

Perfect score: 798

Sequence: 1 atgcagagtctacagctga.....atggaggagctggcaagtga 798

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.2	10.2	1311	4	US-09-068-140A-9
2	81.2	10.2	5150	4	US-09-068-140A-14
3	36.6	4.6	1594	2	US-08-933-750C-61
4	36.6	4.6	1594	2	US-09-234-613-61
5	36	4.5	1643	2	US-08-933-750C-68
6	36	4.5	1643	3	US-09-234-613-68
7	35	4.4	7218	1	US-08-232-463-14
8	32.4	4.1	1777	2	US-08-937-466-5
9	32.4	4.1	1777	3	US-09-172-528-5
10	32.4	4.1	1777	3	US-09-318-199-5
11	32.4	4.1	1777	4	US-09-503-579-5
12	32.4	4.1	1949	2	US-08-937-466-3
13	32.4	4.1	1949	2	US-09-172-528-3
14	32.4	4.1	1949	3	US-09-318-199-3
15	32.4	4.1	1949	4	US-09-503-579-3
16	32.4	4.1	2782	2	US-08-937-466-1
17	32.4	4.1	2782	2	US-09-172-528-1
18	32.4	4.1	2782	3	US-09-318-199-1
19	32.4	4.1	2782	4	US-09-503-579-1
20	31.6	4.0	1816	3	US-09-188-930-262
21	31.2	3.9	4411529	4	US-09-103-840A-1
22	30.8	3.9	1996	1	US-08-385-370-1
23	30.8	3.9	2745	1	US-08-385-370-3
24	30.6	3.8	641	4	US-08-998-416-45
25	30.6	3.8	9633	6	5223423-1
26	30.4	3.8	445	4	US-09-060-756-374
27	30	3.8	2711	4	US-08-891-298-2
28	29.8	3.7	8802	3	US-08-896-449A-1
29	29.8	3.7	8802	3	US-09-132-652-1
30	29.6	3.7	683	2	US-08-997-080-151
31	29.6	3.7	683	2	US-08-997-362-151
32	29.6	3.7	683	4	US-08-095-855-151
33	29.6	3.7	683	4	US-09-324-542-151
34	29.6	3.7	683	4	US-08-630-915A-33
35	29.4	3.7	6246	4	US-08-943-731-640
36	29	3.6	1683	1	US-07-945-283-3
37	29	3.6	1747	1	US-08-202-389-1
38	29	3.6	1817	1	US-08-179-738-1
39	29	3.6	1817	2	US-08-628-145-1
40	29	3.6	1831	6	5215881-1
41	29	3.6	1831	6	5215881-3
42	29	3.6	1833	1	US-08-179-738-6
43	29	3.6	1833	2	US-08-628-145-6
44	29	3.6	1862	1	US-08-179-738-4
45	29	3.6	1862	2	US-08-628-145-4

RESULT 1

US-09-068-140A-9

; Sequence 9, Application US/09068140A

; Patent No. 6281409

; GENERAL INFORMATION:

; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor

; APPLICANT: and Rex Michael Brennan

; TITLE OF INVENTION: Blackcurrant Promoters and Genes

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406-0939

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/068,140A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP96/04807

; FILING DATE: No. 6281409ember 4, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Dinner, Dara L.

; REGISTRATION NUMBER: 33,680

; REFERENCE/DOCKET NUMBER: C70237

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-5017

; TELEFAX: 610-270-5090

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1311 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: CDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Ribes nigrum

; STRAIN: Ben Alder

; US-09-068-140A-9

```
Query Match          10.2%: Score 81.2; DB 4; Length 1311;
Best Local Similarity 50.2%; Pred. No. 1.1e-16;
Matches 274; Conservative 0; Mismatches 248; Indels 24; Gaps 2;

Qy 50 aggcctctgagagattataagaacgagggccctatgagagcccatgaggggctgaacg 109
Db 227 AAGCCCTTGGGTGCGATACTAGAAAGTTGAAGGTCGCCGCGGACTTTACCGTGGCATTTGGT 286

Qy 110 tcacagaacagggcgagggctgccacgccccttatttgcctgtctacgaaagttaa 169
Db 287 CAATGGGTCTCGGTGCAGGACGCTCAGCAGTGTATTTCTCCGTTTACGAGATGTGTA 346

Qy 170 aaagacattgagtgatgaataccacctgggggcaatagccatttgcgaatggtgcg 229
Db 347 AGGAGACTTTTTCATG-----GTGATCCGAGCAATTCGGTCCGACGCCGCTT 397

Qy 230 ccgggtgtgtggcaacattacttcattgatgcagccatgaacccctgcggaagtgtcaagc 289
Db 398 CGGGGTGTTCCGACGCGTGGCAAGCGACGCGGTGATTTACGCCGATGGATGTGTGAAC 457

Qy 290 agagatgcagatgtacaaactcaacataccaccgggtgacagactgtgtacgggcagtgt 349
Db 458 AGAGCTTCAGTTGCAGAGCAGTCCGTACAAGGGTGTGTTGATTGCGTGAGAGGGTGT 517

Qy 350 ggcaaatgaagggcgagggccctttaccgcagctacaccaccagctgaccatgaacg 409
Db 518 TGGTAGAAGAGGATGTCGTCATTTTACGCATCTTATFCGAACAACTGTGTGTCATGAATG 577

Qy 410 ttctttcaagccattcaacttcattgatgcagccatgaacccctgcggaagtgtcaagc 469
Db 578 CCCGCTTACGGCGCTTACCTTCGCCACATATGAAGCACGAAGAGGGTGTGTTGAGG 637

Qy 470 agagcgtgtacaaacccaagctccacgctcct-----ctggagcttgcg 514
Db 638 TGTCCGCGGAGACTGCGAAGCATGAGAAATTTGTTAGTGCATGCTGCTGCTGCTGCTG 697

Qy 515 caggagctgtgtgcgcagccacaaacccactggagcttgcgaacacactgctcaaca 574
Db 698 CTGGAGCTTTGGTGCAGTAGTACCACTCCACTAGATGTTGTCAAAACTCAGTTGCAGT 757

Qy 575 cccagg 580
Db 758 GCCAAG 763

RESULT 2
US-09-068-140A-14
; Sequence 14, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
```

```
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5050
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-14
```

```
Query Match          10.2%: Score 81.2; DB 4; Length 5150;
Best Local Similarity 50.2%; Pred. No. 2.2e-16;
Matches 274; Conservative 0; Mismatches 248; Indels 24; Gaps 2;

Qy 50 aggcctctgagagattataagaacgagggccctatgagagcccatgaggggctgaacg 109
Db 3391 AAGCCCTTGGGTGCGATACTAGAAAGTTGAAGTCCCGCGGACTTTACCGTGGCATTTGGT 3450

Qy 110 tcacagaacagggcgagggccctgccacgccccttatttgcctgtctacgaaagttaa 169
Db 3451 CAATGGGTCTCGGTGCAGGACGCTCAGCAGTGTATTTCTCCGTTTACGAGATGTGTA 3510

Qy 170 aaagacattgagtgatgaataccacctgggggcaatagccatttgcgaatggtgcg 229
Db 3511 AGGAGACTTTTTCATG-----GTGATCCGAGCAATTCGGTCCGACGCCGCTT 3561

Qy 230 ccgggtgtgtggcaacattacttcattgatgcagccatgaacccctgcggaagtgtcaagc 289
Db 3562 CGGGGTGTTCCGACGCGTGGCAAGCGACGCGGTGATTTACGCCGATGGATGTGTGAAC 3621

Qy 290 agagatgcagatgtacaaactcaacataccaccgggtgacagactgtgtacgggcagtgt 349
Db 3622 AGAGTTGCAGTTCCAGAGCAGTCCGTACAAGGGTGTGTTGATTGCGTGAGAGGGTGT 3681

Qy 350 ggcaaatgaagggcgagggccctttaccgcagctacaccaccagctgaccatgaacg 409
Db 3682 TGGTAGAAGAGGATGTCGCGCATTTTACGCATCTTATCGAACAACCTGTGCTCATGAATG 3741

Qy 410 ttctttcaagccattcaacttcattgatgcagccatgaacccctgcggaagtgtcaagc 469
Db 3742 CCCGCTTACGGCGCTTACCTTCGCCACATATGAAGCACGAAGAGGGTGTGTTGAGG 3801

Qy 470 agagcgtgtacaaacccaagctccacgctcct-----ctggagcttgcg 514
Db 3802 TGTCCCGGAGACTGCGCAACGATGAGAAATTTGTTAGTGCATGCTGCTGCTGCTGCTG 3861

Qy 515 caggagctgtgtgcgcagccacaaacccactggagcttgcgaacacactgctcaaca 574
Db 3862 CTGGAGCTTTGGTGCAGTAGTACCACTCCACTAGATGTTGTCAAAACTCAGTTGCAGT 3921

Qy 575 cccagg 580
Db 3922 GCCAAG 3927
```

```
RESULT 3
US-08-933-750C-61
; Sequence 61, Application US/08933750C
; Patent No. 5932442
```

RESULT 4
US-09-234-613-61
; Sequence 61, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti

```

RESULT      5
US-08-933-750C-68
; Sequence 68, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga

```

APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear,
IMMEDIATE SOURCE:
LIBRARY: SYN00AT01
CLONE: 724157
US-08-933-750C-68

Query Match 4.58; Score 36; DB 2; Length 1643;
Best Local Similarity 44.0%; Pred. No. 0.073;
Matches 205; Conservative 0; Mismatches 255; Indels 6; Gaps 1;
QY 26 ctgcccctatcgcaatgtgttgagccctctgagattataagaacgagggcctat 85
DB 482 CTACCCCGCTTACCTGAGCACCATGGATGCTTCGTGAAGATCGTGAGGCACGAGGGCACC 541
QY 86 ggaagcccatgaggggtgaacgtcacagcaacagcgagggcctgcccacgccttt 145
DB 542 GGACCTCTGGAGCGGCTCCGCCACCCCTGGTGATGACTGTGCCAGCTACCGCCATCT 601
QY 146 atttgctgctacgaagaattaaaaagacattgagtgtatccaccctgggggca 205
DB 602 ACTTCACGTGCTTACCACTGACCACTGACCGCCTTCCTGTGTGTGTCGAGCCG-----TGACCT 655
QY 206 atagccatattgccattggtgagcgggtgtgtgcaacattacttcatgacgcca 265
DB 656 CTGACCTTACGACCCATGGTGTGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 715
QY 266 tgaacctcggaagtgtgaagcagaggtatgcagatgtacaactcaccataccaccggg 325
DB 716 TCAGCCCGCTGGAGCTTATGCGGACAAAGCTGACGGCTCAGCATGTGTGCTACCGGAGC 775
QY 326 tgacagactgtgacggcgagtggtgcaaatgaagggcgggcgctttttaccgcagct 385
DB 776 TGGGTGCTCTGTGTCGAACCTGCAGTGCTCAGGGTGGCTGGCGGCTCAGCTGTGGTGGGCT 835

QY 386 acaccaccagctgaccatgaacgttctctttcccaagccattcacttcatgacctatgaat 445
DB 836 GGGGCCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGC 895
QY 446 tctgcagagagcactttaacccccagagacggtacaaacccaagctc 491
DB 896 TGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGGAAGGACCAGACTTC 941
RESULT 6
US-09-234-613-68
Sequence 68, Application US/09234613
Patent No. 6132973
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYN00AT01
CLONE: 724157
US-09-234-613-68

Query Match 4.58; Score 36; DB 3; Length 1643;
Best Local Similarity 44.0%; Pred. No. 0.073;
Matches 205; Conservative 0; Mismatches 255; Indels 6; Gaps 1;
QY 26 ctgcccctatcgcaatgtgttgagccctctgagattataagaacgagggcctat 85
DB 482 CTACCCCGCTTACCTGAGCACCATGGATGCTTCGTGAAGATCGTGAGGCACGAGGGCACC 541
QY 86 ggaagcccatgaggggtgaacgtcacagcaacagcgagggcctgcccacgccttt 145

Db 542 GGACCTCTGGAGCGGCTTCCCGCCACCCCTGGTGATGACTGTGCCAGCTACCGCCATCT 601
Qy 146 attttgctgtctacaaaagttaaagaacattgtagttagtgaatcaaccctgggggca 205
Db 602 ACTTCACCTGCTATGACCAACTGAGGCTTCTCTGTGTGGTGCAGCCC-----TGACCT 555
Qy 206 atagccatattgccaattggtgccccgggtgtgtggaacatttaattatgatgcagcca 265
Db 656 CTGACCTCTACGACCCATGCTGTGGTGGCGCTGGCGCGCTTGGCGCAGCTGACTGTGA 715
Qy 266 tgaacccctgcgaagtgtgcaagcagagagatcagatgtataactcaaccaccggg 325
Db 716 TCAGCCCCCTGAGAGCTTATGCGGAGAAAGCTGACGCTGATGCTGTCACCGGAGC 775
Qy 326 tgacagactgtacggcgagtggtgcaaaatgaaggggcgccgtttttaccgcagct 385
Db 776 TGGGTGCTGTCTTGAAGTGCAGTGGCTCAGGTGGTGGCGCTCACTGTGGTGGGT 835
Qy 386 acaccaccagctgaccatgaacgtttctttccaagccattcaactatgatgcacatgaat 445
Db 836 GGGGCCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGC 895
Qy 446 tcttcagagagcaactttaacccccccagagacggtacaaacccagctc 491
Db 896 TGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGACTTC 941

RESULT 7

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 4.4%; Score 35; DB 1; Length 7218;
Best Local Similarity 4.1%; Pred. No. 0.33;
Matches 14; Conservative 183; Mismatches 148; Indels 0; Gaps 0;

Qy 59 ggaagattataaagaacgagggccctatgagcccatgaggggctgaacgtcacagcaa 118
Db 1383 RRR 1324
Qy 119 caggcgaggggctcccccagcccttattttgctgtctacgaaagttaaaaaagacat 178
Db 1323 RRR 1264
Qy 179 tgagtgtatgaataccaccctgggggcaatagcattatgccaatggtgcggcggtgtg 238
Db 1263 RRR 1204
Qy 239 tggcaacattactcatgagcccatgaaccctgcggaagtgtcaagcagagagatgc 298
Db 1203 RRR 1144
Qy 299 agatgtacaaactcaccataccaccgggtgacagactgtgtacggcgagtggtgcaaaatg 358
Db 1143 RRR 1084
Qy 359 aagggggcgggccctttaccgcagctacacccaccagctgacca 403
Db 1083 RRR 1039

RESULT 8

US-08-937-466-5
; Sequence 5, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

FILING DATE: 08/937,466
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/937,466
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: OSMAN, RICHARD A
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: T97-009
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 343-4341
 TELEFAX: (650) 343-4342
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 1777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-318-199-5

Query Match          4.1%; Score 32.4; DB 3; Length 1777;
Best Local Similarity 44.1%; Pred. No. 1.1;
Matches 135; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 244 acattacttcattgatgcagccatgaacccctgcgggaagtgtcaacgacagaggtgcagatg 303
Db 330 ACTTTTCCCTGGACACCGCCCAAGGTCCGCTGCGAGATCCCAAGGGGAGAAACCCAGGGGCT 389
QY 304 tacaactcaccataccaccgggtgacagactgtgtacgggcagtggtgcaaaatgaaagg 363
Db 390 CAGACGGTGCAGTACCGGGTGTGGGTACCATCCTGACTATGTCGCGACAGAGGCT 449
QY 364 gccggggcctttaccgcagctacacccaccagctgacacccagctgacacgttccttccaaagg 423
Db 450 CCCCAGCGCCCTACAGCGGACTGTGCTGCTGGCTGCCACCCAGATGATGTTTGCCTCC 509
QY 424 attacttcattgaactatgaattcctgcaggagcactttaacccccagagcggtacaac 483
Db 510 ATTCAATTGGCCTTAGACTCTGTCAAGCAGTTCTACACCCCAAGGGAGCGGACCAC 569
QY 484 ccaagctcccagctcctctggaagcttgcgcaggagctgtagctgcgcagccacaacc 543
Db 570 TCCACGGTCGCCATCAGGATTCTGCGAGGCTGCACGACAGGAGCCATGGCAGTGACCTGC 629
QY 544 ccaactg 549
Db 630 GCCCAG 635

RESULT 11
US-09-503-579-5
; Sequence 5, Application US/09503579
; Patent No. 6248561
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,579
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342

;
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-503-579-5

Query Match          4.1%; Score 32.4; DB 4; Length 1777;
Best Local Similarity 44.1%; Pred. No. 1.1;
Matches 135; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 244 acattacttcattgatgcagccatgaacccctgcgggaagtgtcaacgacagaggtgcagatg 303
Db 330 ACTTTTCCCTGGACACCGCCCAAGGTCCGCTGCGAGATCCCAAGGGGAGAAACCCAGGGGCT 389
QY 304 tacaactcaccataccaccgggtgacagactgtgtacgggcagtggtgcaaaatgaaagg 363
Db 390 CAGACGGTGCAGTACCGGGTGTGGGTACCATCCTGACTATGTCGCGACAGAGGCT 449
QY 364 gccggggcctttaccgcagctacacccaccagctgacacccagctgacacgttccttccaaagg 423
Db 450 CCCCAGCGCCCTACAGCGGACTGTGCTGCTGGCTGCCACCCAGATGATGTTTGCCTCC 509
QY 424 attacttcattgaactatgaattcctgcaggagcactttaacccccagagcggtacaac 483
Db 510 ATTCAATTGGCCTTAGACTCTGTCAAGCAGTTCTACACCCCAAGGGAGCGGACCAC 569
QY 484 ccaagctcccagctcctctggaagcttgcgcaggagctgtagctgcgcagccacaacc 543
Db 570 TCCACGGTCGCCATCAGGATTCTGCGAGGCTGCACGACAGGAGCCATGGCAGTGACCTGC 629
QY 544 ccaactg 549
Db 630 GCCCAG 635

RESULT 12
US-08-937-466-3
; Sequence 3, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
```

```

;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1949 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-937-466-3

```

Query Match	4.1%	Score 32.4;	DB 2;	Length 1949;	
Best Local Similarity	44.1%	Pred. No. 1.2;			
Matches 135; Conservative	0;	Mismatches 171;	Indels	0;	Gaps
Qy	244	acattacttc	atgatgcagccatgaacccctgcgaaagtgtgccaagcagagagatgcagatg	303	
Db	330	ACTTTTCCCTGGACACCGCGAAGGTCGCTGTCGAGATCCAAAGGGGAGAACCCAGGGGCT	389		
Qy	304	tacaactcaccataccaccgggtgcagacactgtgtacgggagtcgtggtgcgcaaaatgaaggg	363		
Db	390	CAGACGCTGCAGTACCGGGTGCTGGGTACCATCTGTAGTATGGTGCACAGAGGGT	449		
Qy	364	gcgggggcccctttaccgcagctacaccaccagctgaccatgaacgttcctttcccaagcc	423		
Db	450	CCCCGACGCCCTACAGCGGAGCTGGCTGGCTGCACCCGCCAGATGAGTTTGGCTTCC	509		
Qy	424	attcaacttcagcctatgaattcctgcaggagcacitttaacccccagagacggtacaac	483		
Db	510	ATTGGAATTGGCTCTACGACTCTGTCAAGCAGTTCTACACCCCAAGGGAGCGGACAC	569		
Qy	484	ccaagctccacagcttcctctctggagcttgcgcaggagctgtagctgcccgcagccacaacc	543		
Db	570	TCCAGCTCGCCATCAGGATTCTGGCAGGCTGCACGACAGGAGGCATGGCAGTGACCTGC	629		
Qy	544	ccactg	549		
Db	630	GCCGAG	635		

```

RESULT 13
US-09-172-528-3
; Sequence 3, Application US/09172528
; Patent No. 5952469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/172,528
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341

```

```

; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1949 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-172-528-3

Query Match 4.1%; Score 32.4; DB 2; Length 1949;
Best Local Similarity 44.1%; Pred. No. 1.2;
Matches 135; Conservative 0; Mismatches 171; Indels 0; Gaps

QY 244 acattactcatgatgcagccatgaaccctgcggaagtgtcaagcagaggaatgcagatg 303
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 ACTTTTCCCTGGACACCGCCCAAGTCCGTCTGCAGATCCAAAGGGGAAACCCAGGGGCT 389
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 304 tacaactcacataccacccgggtgacagactgtgtacgggagctgtggcaaaatgaagg 363
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 CAGAGCGTGCAGTACCGCGGTGTGCTGGGTACCATCCTGACTATGGTGCACAGAGGGT 449
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 364 gccggggccctttaccgcagctacaccaccacagctgaccatgaacgttcttccaagc 423
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 CCCGCGACGCCCTACAGCGGACTGGTGCCTGGCTGCACCGCCAGATGAGTTTGGCTCC 509
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 424 attcaactcatgacctatgaattcttcgcagagacactttaacccccagagacgtacaac 483
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 ATTGGAATTGGCTCTACGACTCTGTCAAGCAGTTTACACCCCCAAGGAGCGGACCAC 543
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 484 ccaagctcccaagctcctctctggagcttgcgagaggctgtagctgcccagccacaacc 529
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 TCAGCGTGCCTATCAGGATCTCGCAGGCTGCACGACGAGGCCATGGCAGTGACCTGC 629
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 544 ccactg 549
||| |||
Db 630 GCCCAG 635

RESULT 14
US-09-318-199-3
; Sequence 3, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/318,199
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009

```


Search completed: September 28, 2002, 07:55:54
Job time: 11239 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 06:22:41 ; Search time 4881.77 Seconds
(without alignments)
2206.284 Million cell updates/sec

Title: US-09-870-113-7
Perfect score: 798
Sequence: 1 atgcagagctacagcctga.....agtggaggctggcaagtga 798

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	628.8	78.8	747	10	BI827322
2	587	73.6	826	10	BI254253
3	586	73.4	986	10	BG331197
4	532.2	66.7	576	10	BI057218
5	517	64.8	927	10	BG972351
6	507.4	63.6	908	10	BG697119
7	494	61.9	855	10	BI412175
8	473	59.3	808	10	BG918082
9	472.6	59.2	531	9	AW915209
10	470.6	59.0	563	10	BF193023
11	468	58.6	581	9	AW341177
12	453.6	56.8	830	10	BF143133
13	453.4	56.8	578	10	BM483098
14	437.2	54.8	888	9	AL530804
15	435	54.5	451	10	BF515180
16	433.8	54.4	529	9	AA431276
17	433	54.3	527	9	AI797090

18	432	54.1	526	9	AI797100
19	429.6	53.8	741	9	AL563845
20	427.2	53.5	709	9	AJ397503
21	424.2	53.2	560	10	BI794616
22	423.8	53.1	1131	10	BG295496
23	421	52.8	681	9	AV704087
24	420.2	52.7	645	10	BG964218
25	417.8	52.4	483	9	AA061624
26	405.6	50.8	563	10	BF078769
27	404	50.6	598	10	BE913718
28	398.4	49.9	541	9	AA234031
29	397.6	49.8	574	10	BI185580
30	393.6	49.3	453	10	BF916224
31	391	49.0	488	9	AI797101
32	387.8	48.6	1295	11	AK015770
33	387	48.5	480	10	BI045863
34	387	48.5	718	10	BJ041453
35	384.2	48.1	501	9	AW211366
36	383.2	48.0	423	10	BI339456
37	382.8	48.0	425	10	BI304641
38	381.6	47.8	848	9	AI743110
39	371	46.5	434	10	W53634
40	368.2	46.1	491	10	BF076673
41	363.8	45.6	1209	11	AK006155
42	355.4	44.5	732	9	AI208913
43	350	43.9	466	10	BM258580
44	346.4	43.4	586	10	BG982657
45	332.2	41.6	490	10	BG055067

ALIGNMENTS

RESULT 1
BI827322 747 bp mRNA linear EST 04-OCT-2001
LOCUS 603077989F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169694 5',
DEFINITION mRNA sequence.
ACCESSION BI827322 GI:15938872
VERSION BI827322.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 747)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1422 row: a column: 23
High quality sequence stop: 743.
Location/Qualifiers
1. .747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169694"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library.

BASE COUNT 188 a 206 c 201 g 151 t 1 others
ORIGIN
Query Match 78.8%; Score 628.8; DB 10; Length 747;
Best Local Similarity 99.4%; Pred. No. 3.9e-172;
Matches 652; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 atgcagagctacagctaccagctgcccgcctatcgcaatgtgtgagccctctgg 60
Db 84 ATGCAGAGCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 143
Qy 61 aggattataaagaaacgagggccctatggagggccatgaggggctgaacgtcacagcaaca 120
Db 144 AGGATTATAAAGAACGAGGGCCCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 203
Qy 121 ggcgcagggcctgccacgcctttattttgctgctctacgaaagttaaaaagacattg 180
Db 204 GCGCAGGGCCCTGCCACGCCCTTTATTTCCTGCTACGAAAGTTAAAAAGACATTTG 263
Qy 181 agtgatgtaattccaccctgggggcaatagccattatgccaatgtgctgcccgggtgtg 240
Db 264 AGTGATGTAATCCACCTGGGGGCAATAGCCATATGTCGAATGTGTGGCCGGGTGTG 323
Qy 241 gcaactattctatgctgagccatgaacctgcggagtggtgaagcagaggtgag 300
Db 324 GCAACATTACTTCATGATGATGACCCATGAACCTTGGGAAGTGGTCAAGCAGAGGATGCG 383
Qy 301 atgtacaactaccataccaccgggtgacagactgtgtacgggaagtgtgcaaatgaa 360
Db 384 ATGTACAACCTACCATACCCAGGGGTGACAGACTGTGTACGGGGCAGTGTGGCAAAATGAA 443
Qy 361 ggggcccggggcctttaccgcagctacacacccagctgacctgaacgttctcttccaa 420
Db 444 GGGCCCGGGGCCCTTTTACCGCAGCTACACCCAGCTGACCATGAACGTTCTTTCCAA 503
Qy 421 gccattactctatgacctataaattctctgagggagcactttaacccccagagagcggtac 480
Db 504 GGCATTCACTTCACTGACCTATGAATTCCTGTCAGAGAGCTTTTAAACCCCGCAGAGCGGTAC 563
Qy 481 aaccacaagctccacgtctctctgtgagcttgcgagagctgtgagctgctgagctccacaca 540
Db 564 AACCCAGCTCCACGCTCTCTCTGTGAGCTTCCGAGAGAGCTGTAGCTGCCAGCCACCA 623
Qy 541 acccactggagct-ttgcaaaacactgctcaacacccaggagctctgtgcttgaactc 599
Db 624 GACCCACTGGAGCTTNTTGCAAAACACTGCTCAACACCCAGGAGTCTTGGCTTTTGAAC 683
Qy 600 acacattacagga-catatcacagcagctggttagtgccttcagggcgttatataca 654
Db 684 ACACATTACAGGACCATATACAGGCATGGCTAGTGTGCTTACAGGAGGTATATACAA 739

RESULT 2
B1254253 826 bp mRNA linear EST 17-JUL-2001
LOCUS 602974729F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114190 5',
DEFINITION mRNA sequence.

ACCESSION B1254253
VERSION B1254253.1 GI:14806485
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 826)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L14M11277 row: i column: 07
High quality sequence stop: 825.

FEATURES
source
Location/Qualifiers
1..826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5114190"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."
BASE COUNT 211 a 233 c 208 g 174 t
ORIGIN

Query Match 73.6%; Score 587; DB 10; Length 826;
Best Local Similarity 99.8%; Pred. No. 6.1e-160;
Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 200 gggcaatagccattatcccaatgtgcccgggtgtgtgcaacattactcatgatg 259
Db 1 GGGCAATAGCCATA-TGCCAATGTGTGCGCGGGTGTGTGCAACATTACTTCATGATG 59
Qy 260 cagccatgaacctcggaagtgtcaagcagagatgcagatgtacaactcaccatacc 319
Db 60 CAGCCATGAACCTCGGAAGTGGTCAAGCAGAGATGCAGATGTACAACCTCACCATACC 119
Qy 320 accgggtgacagactgtgtacggcgagtggtgcaaaaagaagggcggtcttttaac 379
Db 120 ACCGGGTGACAGACTGTGTACGGGCAGTGTGCAAAATGAAGGGGCGGGCTTTTAC 179
Qy 380 gcagctacacacccagctgacctgaacgttctcttccaaagccttcaactcatgacct 439
Db 180 GCAGCTACACACCCAGCTGACCATGAAGTTCCTTCCAAAGCCATTCTACTTCATGACCT 239
Qy 440 atgaattcctgcaggagcactttaacccccagagagcggtacaacccaagctccccagtc 499
Db 240 ATGAATTCTTCAGGAGCATTAAACCCCGCAGAGCGGTACAACCCAGCTCCACGTCC 299
Qy 500 tctctggagcttgcaggagctgtagctgcgcagcagcaacacccactggagctttgca 559
Db 300 TCTCTGGAGCTTGGCGCAGGAGCTGTAGCTGCCGCGAGCCACAAACCCCTGAGCTT 359
Qy 560 aaacactcacaacacccagagctcttggcttgaactcacacattacagacatatca 619
Db 360 AAACACTGCTCAACACCCAGGAGTCTTGGCTTTGAACTCACACATTACAGGACATATCA 419
Qy 620 caggcatggttagtgccttcaggagcgttatatacaagtaggtgggtgagccgctatttcc 679
Db 420 CAGGATGCGTAGTGCTTCAGGACGGTATATCAAGTAGGTGGGTGAGCCGCTATTTC 479
Qy 680 gaggggtgcaggccagagtaatttaccagatccccctcacagccatcgcatggtctgtgt 739
Db 480 GAGGGGTGCAGGCCAGAGTAATTTTACCAGATCCCCCTCCACAGCCATCGCATGCTGTGT 539
Qy 740 atgacttctcaaatcacataatcactaaagcagaagagtgaggcgctgcaagtga 798
Db 540 ATGAGTTCTTCAATACCTTAATCACTAAAGCAAGAGAGTGGAGGGCTGCCAACTGA 598
RESULT 3

```

BG331197
LOCUS      602431907F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4549801 5',
DEFINITION mRNA sequence.
ACCESSION BG331197
VERSION   BG331197.1 GI:13137635
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 986)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-r@mail.nih.gov
           Tissue Procurement: DCTD/BTP/Gazdar
           cDNA Library Preparation: Ling Hong/Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L1CM1241 row: m column: 02
           High quality sequence stop: 387.
           Location/Qualifiers
             1..986
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:4549801"
               /clone_lib="NIH_MGC_18"
               /tissue_type="large cell carcinoma"
               /lab_host="DH10B (phage-resistant)"
               /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
               EcoRI; cDNA made by oligo-dT priming. Directionally cloned
               into EcoRI/XhoI sites using the following 5' adaptor:
               GGCACGAG(G). Library constructed by Ling Hong in the
               laboratory of Gerald M. Rubin (University of California,
               Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
               Superscript II RT (Life Technologies). Note: this is a
               NIH_MGC Library."
BASE COUNT 268 a 277 c 248 g 193 t
ORIGIN
Query Match 73.4%; Score 586; DB 10; Length 986;
Best Local Similarity 97.5%; Pred. No. 1.3e-159;
Matches 595; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 187 gtaatccaccctgggggcaatagccatattgccaatggtgcggccgggtgtgtggcaaca 246
Db 2 GTAATCCACCCTGGGGGCAATAGCCATATTGCCAATGTGCGCGGGCTGTGTCGAACA 61

Qy 247 ttacttcgatgatgcagcatgaacctcgcgaagtgttcacagcagagatgcagatgtac 306
Db 62 TTACTTTCATGTATGCAGCATGAACCTCCGGAAGTGTTCACAGAGAGATGCAGATGTAC 121

Qy 307 aactcaccataccaccgggtgcagactgtlacgggcagtggtgcgcaaaatgaagggcc 366
Db 122 AACTCACCATACCACCGGGTGACAGACTGTGTACGGGCGAGTGTGGCAAAATGAAGGGGCC 181

Qy 367 ggggcccctttaccgcagctacacaccaccagctgcacatgaacgttccttccaaagccatt 426
Db 182 GGGGCGCTTTACCGCAGCTACACACCAGCTGACCATGACCATGACGTTCCCTTCCAGCCATT 241

Qy 427 cacttcgatcctatgaattcctgcagagcagcacttaacccccagagcaggtacaccca 486
Db 242 CACTTCATGACCTATGAATTCTCTGCAGGAGCACTTAAACCCCCAGAGCGGTACAACCCA 301

Qy 487 agctcccaagctcctctctgagcttgccagagctgtagctgcgcagagccacaccca 546
Db 302 AGCTCCCAAGCTCCTCTCTGGAGCTTGGCAGGAGCTGTAGCTGCCTGCAGCCACCAACCCA 361

```

```

Qy 547 ctggagcgtttgcaaacacactgctcaacacccaggagtccttgcttgaactcacacatt 606
Db 362 CTGGAGCGTTTGGCAAAACACTGCTCAACACCCAGGAGTCCTTGGCTTGAACTCACACATT 421

Qy 607 acaggacatatcacaggcatggtgctagtccttcaggacagtgatatcaagtagtggtg 666
Db 422 ACAGGACATATCACAGGCATGGCTAGTGCTTTCAGGACGGTATATCAAGTAGTGGGTG 481

Qy 667 accgcctattccgaggggtgcagccagagtaattaccagatccctccacacaccatt 726
Db 482 ACCGCCTATTTCCGACGGGTGCAGGCCAGAGTAATTTACAGATCCCTTCCACAGCCATC 541

Qy 727 gcatgctgtgtatgagttcttcaaatcacctaaatcactaaaggcaagaagtgagg 786
Db 542 GCATGCTGTGTATGAGTTCTTCAATACCTAATCACTAAAGGCAAGACACTGGAGG 601

Qy 787 gctggcaagt 796
Db 602 GCTGGCAAGT 611

RESULT 4
BIOS7218/c
LOCUS      QV3-GN0338-230201-558-a01 GN0338 Homo sapiens cDNA, mRNA sequence.
DEFINITION QV3-GN0338-230201-558-a01 GN0338 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BI057218
VERSION     BI057218.1 GI:14464748
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 576)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
           Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
           Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
           Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
           ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
           Simpson,A.J.
           Shotgun sequencing of the human transcriptome with ORF expressed
           sequence tags
           Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
           20202663
           Contact: Simpson A.J.G.
           Laboratory of Cancer Genetics
           Ludwig Institute for Cancer Research
           Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
           Brazil
           Tel: +55-11-2704922
           Fax: +55-11-2707001
           Email: asimpson@ludwig.org.br
           This sequence was derived from the FAPESP/LICR Human Cancer Genome
           Project. This entry can be seen in the following URL
           (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-GN0338-
           230201-558-a01&t3=2001-02-23&t4=1)
           Seg primer: puc 18 forward
           High quality sequence start: 18
           High quality sequence stop: 575.
           Location/Qualifiers
             1..576
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="GN0338"
               /dev_stage="Adult"
               /note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
               ; Site_2: SmaI; A mini-library was made by cloning
               products derived from ORESTES PCR (U.S. Letters Patent
               application No. 196.716 - Ludwig Institute for Cancer
               Research) profiles into the puc 18 vector. Reverse
               transcription of tissue mRNA and cDNA amplification were
               performed under low stringency conditions."
BASE COUNT 129 a 141 c 162 g 144 t

```


REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10699 row: o column: 10
High quality sequence stop: 878.

FEATURES
source

1..908
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4803681"
/clone_lib="NCI CGAP Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 218 a 236 c 234 g 220 t

ORIGIN

Query Match 63.6%; Score 507.4; DB 10; Length 908;
Best Local Similarity 98.8%; Pred. No. 9.8e-137;
Matches 574; Conservative 0; Mismatches 1; Indels 6; Gaps 6;

QY 223 ggtgcccgggtgtgtgcaacattactatgatgagcagcatgaacccctgcgaagtg 282
|||||
Db 244 ggtgcccgggtgtgtgcaacattactatgatgagcagcatgaacccctgcgaagtg 303
|||||
QY 283 gtcaagcagagatgcagatgtacaactcaccatcacccggtgacagactgtgacgg 342
|||||
Db 304 gtcaagcagagatgcagatgtacaactcaccatcacccggtgacagactgtgacgg 363
|||||
QY 343 gcagtgtgcaaaatgaagggccggggccctttaccgcagctacaccacccagctgacc 402
|||||
Db 364 gcagtgtgcaaaatgaagggccggggccctttaccgcagctacaccacccagctgacc 423
|||||
QY 403 atgaacgttcttccaaagccattcattcatgacctatgaattctcgcagagacattt 462
|||||
Db 424 atgaacgttcttccaaagccattcattcatgacctatgaattctcgcagagacattt 483
|||||
QY 463 a-acccccagagacggtacaaccccagctccacgctcctctctgagcttgcgagagac 521
|||||
Db 484 atacccccagagacggtacaaccccagctccacgctcctctctgagcttgcgagagac 543
|||||
QY 522 tctagctgcgcagccacacccccactgacgcttttgcacaaactgctcaaca-cccagg 580
|||||
Db 544 tctagctgcgcagc-CACAACCCCACTGGACGCTTTGCAAAACACTGCTCAACAGCCAGG 602
|||||
QY 581 agtcttggtcttgactacacaca-ttacagacatatcacaggcattgctagtgccttc 639
|||||
Db 603 agtcttggtcttgactacacaca-ttacagacatatcacaggcattgctagtgccttc 662
|||||
QY 640 a-ggacggttatcaagtaggtgggtgacccctatttccagggggtgcagggccagagt 698
|||||
Db 663 agggagcgtatatacagtaggtgggtgacccgctatttccagggggtgcagggccagagt 722
|||||
QY 699 aatttaccagatccctccacagccatgcagtggtctgtgtgtatgatgtcttcaaatacct 758
|||||
Db 723 aattttaccagatccctccacagccatgcagtggtctgtgtgtatgatgtcttcaaatacct 782
|||||
QY 759 aatcactaaaaaggcaagaag-agtggagggtgctgcaagtga 798

Db 783 AATCACTAAAGGCAAGAAAGTGGAGGCTGGCAAGTGA 823
|||||
RESULT 7
BI412175
LOCUS
DEFINITION
BI412175 855 bp mRNA linear EST 14-AUG-2001
602965377F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5120755 5',
mRNA sequence.
ACCESSION
BI412175
VERSION
BI412175.1 GI:15173098
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11294 row: j column: 20
High quality sequence start: 7
High quality sequence stop: 809.
Location/Qualifiers
1..855
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5120755"
/clone_lib="NCI CGAP Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: p773D-pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGGCCCTCTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 209 a 233 c 241 g 172 t

ORIGIN

Query Match 61.9%; Score 494; DB 10; Length 855;
Best Local Similarity 87.9%; Pred. No. 7.6e-133;
Matches 675; Conservative 0; Mismatches 80; Indels 13; Gaps 12;

QY 1 atgcagagctcacagcctgaccag-ctgccgcctatgcgaatgtgttgaggccctctg 59
|||||
Db 86 atgcagagcctacagcctgaccagcctgctgacgagcctgctgacgagcctgctg 145
|||||
QY 60 gaggattaaagcagggagcctatgagggccctgagggg-ctgagcgtcacagcaa 118
|||||
Db 146 gagatcatcagcagcagaggcctgtggaggccctgctgacgctcagcagcaa 205
|||||
QY 119 caggcgagggcc-tgccacgcctt-ttatttgcctgctacgaaaagtataaaagac 176
|||||
Db 206 caggcgaggccgctgccacgccttgcgtatgttgcctgctacgaaaagtataaaagac 265
|||||
QY 177 a-ttgagtgtatgaatccaccctgggggccaatagccaattgtgcgcgcggt 235

Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
Gene Index
Unpublished (1998)
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information
Seq primer: M13 Reverse.
Location/Qualifiers
1. .531
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RG1CM53"
/clone_lib="Normalized rat embryo, Bento Soares"
/dev_stages="embryo 8, 12, 18 dpc"
/note="Vector: pT7T3Pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT 133 a 157 c 131 g 110 t
ORIGIN
Query Match 59.2%; Score 472.6; DB 9; Length 531;
Best Local Similarity 93.5%; Pred. No. 1.1e-126;
Matches 493; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 241 gcaacattacttcgatgcagccatgaacccctgcggagtggtcaagcagaggtgcag 300
Db 5 GCAAGGTTACTTTCATGTCAGCCATGAATCAGCTGAAGTGGTCAAGCAGCGGATGCAG 64
QY 301 atgtcaactcaccataccacccggtgcagactgtgcaggtgtgcaggtgtgcagaaataca 360
Db 65 ATGTAAACTACCGTACCCGGGTGACAGACTGTGTTCGGGCAGTGTGGCAATGA 124
QY 361 gggggcggggcctttaccgcagctcacaccaccagctgaccatgaacgttctcttccaa 420
Db 125 GGGGCGGGGCGCTTTTACGGCAGCTACACGACCCAGCTGACCATGAATGTTCCCTTCCAA 184
QY 421 gcaattcacttcacactatgaattctcgcaggagacactttaacccccagagcggtac 480
Db 185 GCCATTCACTTTCATACCTATGAGTTCCTGCAGAGCAGCTTTAATCCCGCAGAGCGGTAC 244
QY 481 aacccaagctcccaagctctctgtgagcttcgcgcagagctgtagctgcccagaccaca 540
Db 245 AACCCCGAGCTCCCATGTGCTCTGTGGAGGCTGCGCAGGAGCTGTAGTGCCTGCCGCCACC 304
QY 541 accccactggagctttgcaaaaacactgtctcaacacccaggagctccttgcttgaactca 600
Db 305 ACCCCAGCTGGAGCTTTGCAAAACACTGTCTCAACACCCAGGAGTCCCTGGCCTTGAGCTCA 364
QY 601 cacattacaggacatatcacaggcatgctagtcgcttcagagacggttatatacaagtagt 660
Db 365 AACATTACAGGACACATCACAGGCATGGCTAATGCTTTCAGGACGCTCTATCAAGTAGGC 424
QY 661 ggggtgacgcctatttcgagggtgcagccagagtaatttacagatccctccaca 720
Db 425 GGGGTGACTGCTTACTTCGAGGGGTACAGGCCAGAGTAATTTACAGATCCCTCCCA 484
QY 721 gcatcgcatgctctgtgtatgagttcttcaaatcacttaatacaactaa 767
Db 485 GCCATCGCATGGCTCGGGTATGTAATCTTCAATAACCTAATACACGAA 531
RESULT 10
BF193023
LOCUS BF193023 563 bp mRNA linear EST 02-NOV-2000
DEFINITION 244452 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF193023
VERSION BF193023.1 GI:11076392
KEYWORDS EST.

SOURCE
ORGANISM Sus scrofa
pig.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS 1 (bases 1 to 563)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCAGGAG
Plate: 71 row: A column: 23
Seq primer: ATTTAGGTGCACACTATAG.
Location/Qualifiers
1. .563
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 140 a 174 c 146 g 103 t
ORIGIN
Query Match 59.0%; Score 470.6; DB 10; Length 563;
Best Local Similarity 95.3%; Pred. No. 4.2e-126;
Matches 485; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 290 agagatgcagatgtacaaactcaccatcacccgggtgcagactgtgtacggcgagtg 349
Db 1 AGAGGATGCAGATGAACAACCTACCCCTACCCGCGGTGACAGACTGTGTACGGGCGAGTGT 60
QY 350 ggcataaatgaaggggcgcgggcctttttaccgcagctcacaccaccagctgaccatgaacg 409
Db 61 GGCAAAACGAAGGGCGCGGGCGCTTTTACCGCAGCTACACCACCCAGCTCAACCATGAACG 120
QY 410 ttcttcttcaagcattcacttcacatgcacatgaattcctcagcagacacttaaccccc 469
Db 121 TTCCTCTCCAGGCCATCCACTTCATGACCTATGAATTCCTGCAGGAGCACTTTAACCCCC 180
QY 470 agagacggtacaaaccccaagctccctctctgtgagcttgcgcagagctgtagctg 529
Db 181 AGAGACGTTACAAACCCAGCTCCACGTCCTCTCGGAGGCTGCGCAGAGCTGTAGCTG 240
QY 530 ccgcagcacacaccccaactggacgtttgcaaaacactgctcaacacccaggagctccttgg 589
Db 241 CGGCGCGCGCAACCCCACTGGAGGTTTGCAAAACACTGCTCAATACCCAGGAGTCCCTGG 300
QY 590 ctttgaactcaacatttacaggacacatacaacaggcatggttagtccttcaggagcggtat 649
Db 301 CTTTGAACTCAAACATTACAGGACACATCACAGGCATGGCTAGTGCCTTCAGGACGGTGT 360
QY 650 atcaagtaggtgaggtgacccctatttccagggttcagggccagcagagtaattaccaca 709
Db 361 ATCAAGTAGGGGAGTGAACCGCTTACTTCCGAGGGGTGCAGGCGAGAGTAATTTACACAGA 420
QY 710 tcccttcacacagccatcgctggtctgtgtatgagttcttcaaatcacctaatacaaaaa 769
|||||

```

Db 421 TCCCTCCACGCCATCGGTGTGTATGAGTCTTCAATAACCTGATCACTAAAC 480
Qy 770 gcaagaagtgaggctggcgaagtga 798
Db 481 GCAAGAGAGTGGAGGCGCAGCAAGTCA 509

RESULT 11
AW341177
LOCUS AW341177 581 bp mRNA linear EST 31-JAN-2000
DEFINITION x295e05.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2871968 3'
similar to TR:Q94638 Q94638 MITOCHONDRIAL SOLUTE CARRIER. ;, mRNA
sequence.
ACCESSION AW341177
VERSION AW341177.1 GI:6837803
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco
High quality sequence stop: 369.
FEATURES
Location/Qualifiers
source 1..581
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2871968"
/clone_lib="NCI_CGAP_Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI_CGAP_Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (clones
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 138 a 172 c 135 g 134 t 2 others
ORIGIN

Query Match 58.6%; Score 468; DB 9; Length 581;
Best Local Similarity 97.5%; Pred. No. 2.4e-125;
Matches 474; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 279 agtggtcaagcagagatgcagatgtacaactaccataccacgggtgacagactgtgt 338
Db 95 AGTGTCAAGCAGAGAGTGCAGATGTACAACCTACCATTACCCCGGGTGACAGACTGTGT 154
Qy 339 acgggagctgtggcaaatgaagggccggccctttaccgcagctacacacccagct 398
Db 155 ACGGCGACTGTGGCAAAATGAAGGGCGGGCCCTTTTACCGCAGCTACACCCCGAGCT 214
Qy 399 gaccatgaacgttctcttcccaagccattcacttcacgtatgatgaattcctgcaggagca 458

```

```

Db 215 GACCATGAACGTTCTCTTCCAAAGCCATTCACTTCATGACCTATGAATTCCTGCAGGAGCA 274
Qy 459 ctttaacccccagagagcgggtacaacccaagctccacgctctctctctgagcttgcgagg 518
Db 275 CTTTAAACCCCCAGAGAGCGGTACAACCCAAAGCTCCACAGTCTCTCTGGAGTTCGCGAGG 334
Qy 519 agctgtagctgcgcagccacaacccccactgtgacgtttgcaaaacactgctcaacaccca 578
Db 335 AGCTGTAGTCCGCGCACCAACACCCACCTGGACGTTTGGCAAAACACTGCTCAACACCCA 394
Qy 579 ggaagctcttggttgaactcacacattacaggagcatatcacaggcatggctagtcctt 638
Db 395 GGAGTCTCTGGCTTTGAACATCACACATTCAGAGACATATCACAGGCATGGCTAGTGCTT 454
Qy 639 caggacggtatatcaagtagtggtgggtgacgcctatttcgaggggtgcagggcagagt 698
Db 455 CAGGACCGTATATCAAGTAGTGGGTGACGCCCTATTTTCGAAGGGGTGCAGGCCAGAGT 514
Qy 699 aattaccagatccctccacagccatcgcatggtctgtgtatgagttcttcaataacct 758
Db 515 AAATTACCAGATCCCTCCACAGCCATCGCATGGTCTGTATGAGTTCCTCCANATACCT 574
Qy 759 aatcac 764
Db 575 AATCAC 580

RESULT 12
BF143133
LOCUS BF143133
DEFINITION 601788025F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4015758 5',
mRNA sequence.
ACCESSION BF143133
VERSION BF143133.1 GI:10982173
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 830)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9262 row: a column: 07
High quality sequence stop: 607.
FEATURES
Location/Qualifiers
source 1..830
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4015758"
/clone_lib="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 202 a 239 c 220 g 169 t
ORIGIN

```

Query Match 56.8%; Score 453.6; DB 10; Length 830;
Best Local Similarity 92.4%; Pred. No. 4.4e-121;
Matches 477; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 283 gtcaagcagagatcagatgtacaactcacaccatcacaccgggtgacagactgtgtacgg 342
DB 1 gtcaagcagagatcagatgtacaactcacaccatcacaccgggtgacagactgtgtacgg 60

QY 343 gcaagtgtggcaaaatgaaggccggggccttttaccgcagctacacaccagctgacc 402
DB 61 gcagtggtggcaaaatgaaggccggggccttttaccgcagctacacaccagctgacc 120

QY 403 atgaacgttccttccaaagccattcactatcatgacctatgaattcctcagagacacttt 462
DB 121 atgaatgtcccccttccaaagccattcactatcatgacctatgaattcctcagagacacttt 180

QY 463 aacccccagagcgtacaacccaaagctcccaagctcctctgtgagctgcagagact 522
DB 181 aacccccagagcgtacaacccaaagctcctctgtgagctgcagagact 240

QY 523 gtactgtccgcagcagcaacccccactggactgtttgcacaaactgtctcaacaccagag 582
DB 241 gtactgtccgcagcagcaacccccactggactgtttgcacaaactgtctcaacaccagga 300

QY 583 tccctgtgttgaaactcacacattacagagacatatcagggcattgcttagtcctcagg 642
DB 301 tccctgtgtgttgaaactcacacattacagagacatatcagggcattgcttagtcctcagg 360

QY 643 acggtatatcaagtgtgggtgacccctatttccaggggtgcagggcagagtaatt 702
DB 361 acggtgtatcaagtgtgggtgacccctatttccaggggtgcagggcagagtaatt 420

QY 703 taccagatccccctccacagcagcagctgtgtgtgtatgagttctcaaatcacctaatt 762
DB 421 taccagatccccctccacagcagcagctgtgtgtgtatgagttctcaaatcacctaatt 480

QY 763 actaaaggcgaagaagtgagggtggcgaagtga 798
DB 481 actaaaggcgaagaagtgagggtggcgaagtga 516

RESULT 13
BM483098 578 bp mRNA linear EST 05-FEB-2002
LOCUS 536233 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BM483098
ACCESSION BM483098
VERSION BM483098.1 GI:18533417
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 578)
Smith, T.P.L., Grosche, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
Perte, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers

FORWARD: AGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGCTACGACG
Plate: 5 row: F column: 16
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1..578
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT 139 a 184 c 141 g 114 t
ORIGIN
Query Match 56.8%; Score 453.4; DB 10; Length 578;
Best Local Similarity 94.7%; Pred. No. 4.4e-121;
Matches 469; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 304 tacaactcacatcacaccgggtgacagactgtgtacgggcagtggtgcaaaatgaagg 363
DB 1 TACAACCTCACCGTACCACCGGGGTGACAGACTGTGTCCGGGCACTGTGGCAGAACGAAGGG 60

QY 364 gccggggcctttaccgcagctacacaccagctgcacatgaacgttcccttccaaagcc 423
DB 61 GCTGGGGCCTTTTACCAGGCTACACCCAGCTCACCAGCTCACCAGTTCCTTCCAAAGCC 120

QY 424 attacttcacatgatgaattcctgcagagcactttaaccccgagacggtacacac 483
DB 121 ATTCACTTCATGACCTACGAATTCCTGCAGGAGCACTTTAACCCCGACAGACGGGTACAC 180

QY 484 ccaagctccccagctcctctctgtggagctgtgcagagagctgtagctgcgcgagccacacac 543
DB 181 CCCAGCTCCACGCTCTCTCCGGAGCCTGTGCAGGAGCTGTAGCTGCCGCTGCCACAACC 240

QY 544 ccactggagcttgcacaaactgctcaacccagggagctcctgtgttgaactcacac 603
DB 241 CCACCTGGAGCTTGCACAAACACTGCTCAACACCCAGGAAATCCCTGGCTTTGAACCTCAAC 300

QY 604 attacagacatatcacaggcatggctgtgagctgtgcagagagctgtagctgcgcgagccacacac 663
DB 301 CTTACAGGACATATCACAGGATGGCTAGTGCCTTCAGSAGCGGTGTATCAAGTAGGCGGG 360

QY 664 gtgaccgcctatttccgaggggtgcagggcagagtaatttaccagatccccctccacagcc 723
DB 361 GTGACTGCCTACTTCCGAGGGGTGCAGGCTAGAGTCAITTTACCAGATCCCTCCACAGCC 420

QY 724 atcgcatggtctgtgtatgagttcttcaaaatcacctaatacctaaagcagaagagtg 783
DB 421 ATCGCTGGCTGTGTATGAGTTCTTCAAAATACCTGTACTATAAAGGCAAGAGAGTGG 480

QY 784 agggctggcgaagtga 798
DB 481 AGGCAGGCAAGTGA 495

RESULT 14
AL530804
LOCUS AL530804 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD008YH17 5
DEFINITION AL530804 prime, mRNA sequence.
ACCESSION AL530804
VERSION AL530804.1 GI:12794297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 888)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1. .888
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DD008YH17"
 /clone_lib="LTI_NFL001_NBC4"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 168 a 244 c 322 g 139 t 15 others
 ORIGIN

Query Match 54.9%; Score 437.2; DB 9; Length 888;
 Best Local Similarity 97.2%; Pred. No. 2.7e-116;
 Matches 459; Conservative 6; Mismatches 5; Indels 2; Gaps 2;

Qy 1 atgcagagtacagctaccagctgcccgtatccgctatcgcaatgtgttgaggccctctg 60
 Db 417 ATGCAGAGCTACAGCTGACCCAGCTGCCCCGTATCGCAATGTTGGAGGCCCTCTG 476

Qy 61 aggattataagaacggggcctatgagggccatgagggggcgtgaacgtcacagcaaca 120
 Db 477 AGGATTATAAGAACGGGGCCTATGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 536

Qy 121 ggcgagggcctgccacgccttatttgctgtctacgaaagttaaaagacattg 180
 Db 537 GCGCAGGGCTGCCACGCCCTTATTTCCTGTCBACGAAAAAGTTAAAAAGACATTG 596

Qy 181 agtatgaatcaaccctggggcgaatagccatttcccaatggtgcgcgggtgtgtg 240
 Db 597 AGTATGTTATCCACCTGGGGGCAATAGCCATATTCCAATGTCGGCGGGTGTGTG 656

Qy 241 gcaacattacttcgatgcagccatgaacctgcggaagtgttcaagcagagagtcag 300
 Db 657 GCAACATTACTTCATGATGACGACCATGACCCCTCGGAAGTGTGTCAGSAGAGGATGCAG 716

Qy 301 atgtacaactcaaccataccaccgggtgacagactgtgtacgggcagtggt-ggcaaaatga 359
 Db 717 ATGTACAACCTACCATACCACCGGGTGACAGACTGTGTACGGGCAGTGTGGCAAAATGA 776

Qy 360 agggcgccggccctttacgcagcactacaccaccactgaccatgaactctctcca 419
 Db 777 AGGGCGCGGGCCCTTTACCGCAGCTACACACCCCAATGACGACGTCGACGTCCTCTTCCA 836

Qy 420 agccattcaactca-tgacctatgaattcctgcaggagcactttaaccccca 470
 Db 837 AGCCATTACTTCAWTGACCTATGAATCTCTGCAGGAGCCCTTTAACCCCCA 888

RESULT 15
 BF515180 451 bp mRNA linear EST 07-DEC-2000
 LOCUS UI-H-BW1-anl-b-10-0-UI.sl NCI_CGAP_Sub7 Homo sapiens cDNA clone
 DEFINITION IMAGE:3082698 3', mRNA sequence.
 ACCESSION BF515180

VERSION BF515180.1 GI:11600447
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1. (bases 1 to 451)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 NCI-CGAP clone distribution information can be found through the
 I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbrrp/image/image.html
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES
 source
 1. .451
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3082698"
 /clone_lib="NCI_CGAP_Sub7"
 /lab_host="DH10B (Life Technologies)"
 /notes="Vector: pT7T3D-Pac (Pharmacia).with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; NCI_CGAP_Sub7
 is a subcloned library derived from NCI_CGAP_Sub6. The
 NCI_CGAP_Sub7 library had 12 million recombinants. A
 single-stranded DNA preparation of NCI_CGAP_Sub6 was used
 as a tracer in a subtractive hybridization with a driver
 comprising: the IMAGE pool (NCI_CGAP_Kid3 pool 1 LLAM
 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs
 132376-132391), 1456008-1456775, 1500552-1502855);
 NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778
 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903,
 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582,
 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439
); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720,
 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983
 , 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459,
 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
 1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1
 LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255
 , 1144584-1145351). (6% of the driver population), plus a
 pool of 3,840 arrayed clones from NCI_CGAP_Sub1 (IMAGE
 CloneIDs 2708616-2710535) and NCI_CGAP_Sub2 (IMAGE
 CloneIDs 2710536-2712455) (4% of the driver population
), plus a pool of 11,136 clones from NCI_CGAP_Sub3 (IMAGE
 CloneIDs 2712456-2723591) (10% of the driver population),
 plus a pool of 5,472 clones from NCI_CGAP_Sub4 (IMAGE
 CloneIDs 2723592-2729326) (40% of the driver population),
 plus a pool of 4032 clones from NCI_CGAP_Sub6 (IMAGE
 CloneIDs 2728969-2733190) (40% of the driver population).
 Subtraction was performed as previously described (Bonaldi
 , Lennon & Soares (1996): Normalization and Subtraction:
 Two Approaches To Facilitate Gene Discovery. Genome
 Research 6, 791-806.
 TAG_LIB=NCI_CGAP-Lu13
 TAG_TISSUE=lung
 TAG_SEQ=GCGCG"

BASE COUNT 112 a 129 c 109 g 101 t
 ORIGIN

Query Match 54.5%; Score 435; DB 10; Length 451;
 Best Local Similarity 100.0%; Pred. No. 9e-116;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	223	ggtcgccggctgtgtgccaattactctcatgtatcgaccatgaacctgcgaagt	282
Db	17	GGTCGGCCGGGTGTGTGGCAACATTACTTTCATGATGCACCATGAACCTCGGAAGTG	76
Qy	283	gtcaagcagagatgcagatgtacaactcaacataccaccgagtacacagtatacag	342
Db	77	GTCNAGCAGAGATGCAGATGTACACTCACCATACCCGGGTGACACACTGTGTACGG	136
Qy	343	gcagttgtgcaaatgaaggggccggggccctttaccgcagctacaccaccagctgacc	402
Db	137	GCAGTGTGGCAAAATGAAGGGCCGGGGCCCTTTACCGCAGCTACACCAACCCAGCTGACC	196
Qy	403	atgaagcttccttccaagccattcaacttcaatgaacctatgaattcctgcagagacatt	462
Db	197	ATGAACGTTCTTTCCAAGGCATTCTACTTCATGACCTATGAATTCCTGCAGGAGCACTTT	256
Qy	463	aacccccagagacggtacaaacccaagctcccagctcctctctggagcttgcgcagagct	522
Db	257	AACCCCCAGAGACGGTACACCCCAAGCTCCCAGTCTCTCTGGAGCTTGGCAGGAGCT	316
Qy	523	gtagctgccgagccaaacccccacttggagttgcaaacactgctcaacccccaggag	582
Db	317	GTAGCTGCCGAGCCAAACCCCACTGGAGTTTGTGAAAACACTGCTCAACACCCAGGAG	376
Qy	583	tccttggcttgaactcacacattacagacatatcacaggatggctagtccttcagg	642
Db	377	TCCTTGGCTTTGAACCTCACACATTACAGGACATATCACAGGCATGCTAGTGCCCTTCAGG	436
Qy	643	acggatatcaagta	657
Db	437	ACGGTATATCAAGTA	451

Search completed: September 28, 2002, 06:22:48
Job time: 9189 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:57:11 ; Search time 5287.54 Seconds
(without alignments)
1127.947 Million cell updates/sec

Title: US-09-870-113-9
Perfect score: 285
Sequence: 1 atgcagagtctacagctga.....tgaacctgcggaaggtcta 285

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.ot.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	280	98.2	1322	6	AX061229	Sequence
2	280	98.2	1448	9	AF327402	Homo sapi
3	280	98.2	1448	17	HS303077	Homo sapi
4	273	95.8	401	6	AX071440	Sequence
5	225	78.9	1889	9	AF327403	Homo sapi
6	225	78.9	1889	17	HS303078	Homo sapi
7	225	78.9	123160	9	AL353719	Human DNA
8	202.6	71.1	221062	2	AC096351	Rattus no
9	139.2	48.8	4112	10	AF288621	Mus muscu
10	128	44.9	1429	9	AF223466	Homo sapi
11	125.2	43.9	163205	2	AC099368	Rattus no
12	109.8	38.5	5869	10	AF361699	Mus muscu
13	109.8	38.5	191702	2	AC108878	Mus muscu
14	105.2	36.9	483	6	AX306258	Sequence
15	105	36.8	108765	2	AF216674	Homo sapi
16	103.4	36.3	858	9	BC015013	Homo sapi
17	103.4	36.3	3415	9	AY032628	Homo sapi
18	103.4	36.3	174445	9	AC051642	Homo sapi
19	103.4	36.3	182504	2	AC022597	Homo sapi
20	51	17.9	30891	2	AC100095	Mus muscu
21	50.4	17.7	1072	10	BC023172	Mus muscu
22	49.6	17.4	34305	3	CEW02812	Caenorhabdi
23	42.2	14.8	3037	3	AY060268	Drosophil
24	40.6	14.2	2878	3	AF217402	Drosophil
25	36.2	12.7	1243	3	OGU45997	Onchocerca
26	36.2	12.7	1281	3	OVU45998	Onchocerca
27	36	12.6	111464	8	F10K1	Sequence
28	35.6	12.5	180397	2	AF003615	Oryza sat
29	35.2	12.4	90935	2	AP003917	Oryza sat
30	35.2	12.4	169672	2	AC103286	Rattus no
31	34.6	12.1	223538	2	AC087150	Mus muscu
32	34.4	12.1	1517	9	AF125531	Homo sapi
33	34.2	12.0	28993	3	LMFL5808	Leishmani
34	34	11.9	657	8	AF121142	Gracilari
35	34	11.9	7242	8	MITARRNG	T.aestivum
36	34	11.9	103787	9	AC093659	Homo sapi
37	34	11.9	152248	2	AC027522	Homo sapi
38	34	11.9	184034	2	AC024156	Homo sapi
39	33.8	11.9	1154	4	AF127029	Bos tauru
40	33.8	11.9	64790	2	AC099356	Rattus no
41	33.8	11.9	125020	9	AF429315	Homo sapi
42	33.8	11.9	174591	2	AC103093	Rattus no
43	33.6	11.8	1366	8	AY063977	Arabidops
44	33.6	11.8	58856	8	AC004165	Arabidops
45	33	11.6	114642	9	AL357033	Human DNA

ALIGNMENTS

RESULT 1
LOCUS AX061229
DEFINITION Sequence 76 from Patent WO0078953.
ACCESSION AX061229
VERSION AX061229.1 GI:12406365
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1322)
AUTHORS Lal, P., Yang, J., Yue, H., Hillman, J.L., Tang, Y.T., Bandman, O., Burford, N., Baughn, M.R., Azimzai, Y., Lu, D.A., Au-Young, J. and Patterson, C.
TITLE Human transport proteins
JOURNAL Patent: WO 0078953-A 76 28-DEC-2000;
FEATURES Location/Qualifiers
1..1322

BASE COUNT 328 a 376 c 351 g 267 t

ORIGIN

Query Match 98.2%; Score 280; DB 6; Length 1322;
Best Local Similarity 100.0%; Pred. No. 6.5e-83;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 atgcagagctcacagcctgacccagctgcccgcctatgcgaatgtgtgagggccctctgg 60
|||||
Db 186 ATGAGAGCTACAGCCTGACCCAGCTGCCGCTATGCCATGTGTTGAGGCCCTCTGG 245
|||||

Oy 61 aggattataagaacaggagggcctatggagggccatgaggggctgaaacgtcacagcaaca 120
|||||

Db 246 AGGATTATAAGAACGAGGGCGCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 305
|||||

Oy 121 ggcgagggcctgcccagcccttatttttgcctgctacgaaagttaaaaaagacattg 180
|||||

Db 306 GCGCAGGGCCTGCCACGCCCTTTATTTTCCTGCTACGAAAAAGTTAAAAAGACATTG 365
|||||

Oy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatggtgcggccgggtgtgtg 240
|||||

Db 366 AGTGATGTAAATCCACCCCTGGGGCAATAGCCATATTGCCAATGGTGCGCCGGGTGTGTG 425
|||||

Oy 241 gcaacattactctatgcagcgcctgacccctgcggaag 280
|||||

Db 426 GCAACATTACTTATGATGATGAGCCATGAACCTCGCGGAAG 465
|||||

RESULT 2

AF327402 AF327402 1448 bp mRNA linear PRI 02-MAY-2001
LOCUS Homo sapiens putative mitochondrial solute carrier splice variant
DEFINITION mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.

ACCESSION AF327402

VERSION AF327402.1 GI:13926046

KEYWORDS

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1448)

Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B.,

Schweyen, R., Larsson, C. and Suomalainen, A.

Characterization of a novel human putative mitochondrial

transporter homologous to the yeast mitochondrial RNA splicing

proteins 3 and 4

FEBS Lett. 494 (1-2), 79-84 (2001)

JOURNAL MEDLINE

PUBMED

11297739

2 (bases 1 to 1448)

Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B., Schweyen, R.,

Larsson, C. and Suomalainen, A.

Direct Submission

Submitted (08-DEC-2000) Human Molecular Genetics, National Public

Health Institute, Mannerheimintie 166, Helsinki 00300, Finland

FEATURES

Source

1. .1448

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="10"

30. .1124

/note="similar to mitochondrial RNA splicing protein 3/4;

alternatively spliced"

/codon_start=1

/product="putative mitochondrial solute carrier splice

variant"

/protein_id="AAK49519.1"

/db_xref="GI:13926047"

/translation="MELEGRAGGVAGGPAAGRSPGSEALLDGNLQGVGRGAGGG

EAGCRPPVQDDSPDYPALPAGATVTHMVAGAVGLEHCVMYPIDCVKTRMOS

LQDPAPARYNLEALWRIIRTEGLRPMRGLNVTATGAGPAHALYFACYELKRLTSL

DVIPHGNSHIANGAGCVATLLHDAAMNPAEVVKQRMQYNSPVRHVTDCVRAVWON
EGAGAFYVYTTLTMMVFOATHFMYEFLQEHFNPPORRYNPSSHLSCAGAGVAA
AATTPLDVCCKTLTNTQESLAINSHITGHITGMASAFRTVYQGVGTAYFRGVQARVIY
QIPSTAINSVYEFFKYLITKROEWRAG"

BASE COUNT 322 a 408 c 435 g 283 t
ORIGIN

Query Match 98.2%; Score 280; DB 9; Length 1448;

Best Local Similarity 100.0%; Pred. No. 6.6e-83;

Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 atgcagagctcacagcctgacccagctgcccgcctatgcgaatgtgtgagggccctctgg 60

Db 327 ATGAGAGCTACAGCCTGACCCAGCTGCCGCTATGCCAATGTGTTGAGGCCCTCTGG 386

Oy 61 aggattataagaacaggagggcctatggagggccatgaggggctgaaacgtcacagcaaca 120

Db 387 AGGATTATAAGAACGAGGGCGCTATGAGGGCCCATGAGGGGCTGAACGTCACAGCAACA 446

Oy 121 ggcgagggcctgcccagcccttatttttgcctgctacgaaagttaaaaaagacattg 180

Db 447 GCGCAGGGCCTGCCACGCCCTTTATTTTCCTGCTACGAAAAAGTTAAAAAGACATTG 506

Oy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatggtgcggccgggtgtgtg 240

Db 507 AGTGATGTAAATCCACCCCTGGGGCAATAGCCATATTGCCAATGGTGCGCCGGGTGTGTG 566

Oy 241 gcaacattactctatgcagcgcctgacccctgcggaag 280

Db 567 GCAACATTACTTATGATGATGAGCCATGAACCTCGCGGAAG 606

RESULT 3

HSA303077

ID HSA303077 standard; RNA; HUM; 1448 BP.

XX AC AJ303077;

XX SV AJ303077.1

XX DT 02-FEB-2001 (Rel. 66, Created)

XX DT 26-APR-2001 (Rel. 67, Last updated, Version 2)

XX DE Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMRS3/4

gene), 1448 bp cDNA splice variant

XX KW HMRS3/4 gene; mitochondrial RNA splicing protein 3/4.

XX OS Homo sapiens (human)

XX OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

XX OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX RN [1]

RN RP 1-1448

RA Nikali K.;

RT ;

RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.

RL Nikali K., Human Molecular Genetics, National Public Health Institute,

RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.

XX RN [2]

RA Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,

RA Larsson C., Suomalainen A.;

RT "Characterization of a novel human putative mitochondrial transporter

homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";

FEBS Lett. 494:79-84(2001).

XX Key

Location/Qualifiers

1. .1448

/db_xref="taxon:9606"

/organism="Homo sapiens"

FT CDS 30..1124
/gene="HMR3/4"
/product="mitochondrial RNA splicing protein 3/4"
/function="putative mitochondrial solute carrier"
/protein_id="CAC27996.1"
/translation="MELEGRGAGVAGGPAAGPGRSPGSEALLDGLWLGVRGAGGGE
AGACRPVRQDPDGPDPYATVTHVAVAGALEHCYVMPIDCVKTRQSLQ
PDPAARNVLEALWRIIRTEGLRWRPRLNVTATGAPAHALYFACYEKLKLTLSQVI
HPGNSHLANGAACCVATLLHDAAMNPAEVVKQRMQNSPYHRVTCVRAVWONEGAG
AFYRSYTLQTMNVPFOAIHMTVEFLQEHENPQRRYNPSSHVLSGACAGAAAAATTP
LDVCKTLNTOESLALNSHJTGHTGHASAFRTYQVGGVTAYFRGVQARVIYQIPSTA
IANSVYEFFKLIITKQEEWRAGK"
XX
SQ Sequence 1448 BP; 322 A; 408 C; 435 G; 283 T; 0 other;

Query Match 98.2%; Score 280; DB 17; Length 1448;
Best Local Similarity 100.0%; Pred. No. 6.6e-83;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
Db 327 ATGCAGAGTCTACAGCTGACCGCTGCGCGCTATCGCAATGTGTGGAGGCCCTCTGG 386

Qy 61 aggattataagaacggagggccctatgaggccccatgaggggctgaacgtcacagcaaca 120
|||||
Db 387 AGGATTATAAGAACGAGGGCCCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 446

Qy 121 ggcgcagggcctgcccagccctttatttgcctactacgaaagttaaaaagacattg 180
|||||
Db 447 GCGCAGGGCCTGCGCCACGCCCTTTATTTCCTGCTACGAAAGTTAAAAAGACATTG 506

Qy 181 agtgatgaatccaccctggggccaatgccaattggtcgccgggtgtgtg 240
|||||
Db 507 AGTGATGTAATCCACCCTGGGGCAATAGCCATATGCCAATGTGTGGCGGGTGTGTG 566

Qy 241 gcaacattactcatgatgcagccatgaacctgcggaag 280
|||||
Db 567 GCAACATTACTTCATGATGCAGCCATGAACCCCTGCGGAAG 606

RESULT 4
AX071440 401 bp DNA linear PAT 25-JAN-2001
LOCUS
DEFINITION Sequence 1912 from Patent WO0102568.
ACCESSION AX071440
VERSION AX071440.1 GI:12581791
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 401)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Lanson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 1912 11-JAN-2001;
CHIRON CORPORATION (US); HYSEQ, INC. (US)
FEATURES
source
1..401
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1..401
/note="n = A,T,C or G"
BASE COUNT 100 a 104 c 120 g 76 t 1 others
ORIGIN

Query Match 95.8%; Score 273; DB 6; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.2e-80;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 gtctacagctgacccagctgccgctatcgcaatgtgttgaggccctctgagagatta 67
|||||
Db 9 GTCACAGCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTCGAGATTA 68
|||||
Qy 68 taagaacggaggccctatgagggcccatgaggggctgaacgtcacagcaacagcgag 127
|||||
Db 69 TAAGAAGGAGGGCCCTATGAGGGCCCATGAGGGGCTGAACGTCACAGCAACAGGCGAG 128
|||||
Qy 128 ggcctgcccagcccttatttgcctactacgaaagttaaaaagacattgaagtgtg 187
|||||
Db 129 GGCCTGCCACGCCCTTTATTTCCTGCTACGAAAGTTAAAAAGACATTGAGTGATG 188
|||||
Qy 188 taatcacccctggggccaatgccaatttgcgaatgggtgcgcccgggtgtgtggcaacat 247
|||||
Db 189 TAATCACCCCTGGGGCAATAGCCATATTGCCAATGTGTGGCGCGGTGTGTGGCAACAT 248
|||||
Qy 248 tacttcagtgcagcccatgaacctgcggaag 280
|||||
Db 249 TACTTCATGATGCAGCCCATGAACCCCTGCGGAAG 281

RESULT 5
AF327403 1889 bp mRNA linear PRI 02-MAY-2001
LOCUS
DEFINITION Homo sapiens putative mitochondrial solute carrier splice variant
mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.
ACCESSION AF327403
VERSION AF327403.1 GI:13926049
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1889)
AUTHORS Li,F.Y., Nikali,K., Gregan,J., Leibiger,I., Leibiger,B.,
Schweyen,R., Larsson,C. and Suomalainen,A.
TITLE Characterization of a novel human putative mitochondrial
transporter homologous to the yeast mitochondrial RNA splicing
proteins 3 and 4
JOURNAL FEBS Lett. 494 (1-2), 79-84 (2001)
MEDLINE 21195335
PUBMED 11297739
REFERENCE 2 (bases 1 to 1889)
AUTHORS Li,F., Nikali,K., Gregan,J., Leibiger,I., Leibiger,B., Schweyen,R.,
Larsson,C. and Suomalainen,A.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-2000) Human Molecular Genetics, National Public
Health Institute, Mennerheimintie 166, Helsinki 00300, Finland
FEATURES
Location/Qualifiers
source
1..1889
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q24"
1032..1565
/note="similar to mitochondrial RNA splicing protein 3/4;
alternatively spliced"
/codon_start=1
/product="putative mitochondrial solute carrier splice
variant"
/protein_id="AAK49520.1"
/db_xref="GI:13926050"

BASE COUNT 404 a 508 c 544 g 433 t
ORIGIN
/translation="MNPAEVVKQRMQNSPYHRVTCVRAVWONEGAGAFYRSYTTQ
LTMNVPFOAIHMTVEFLQEHENPQRRYNPSSHVLSGACAGAAAAATPDDVCKTL
NTOESLALNSHJTGHTGHASAFRTYQVGGVTAYFRGVQARVIYQIPSTAIANSVYE
FFKYLITKQEEWRAGK"

Query Match 78.9%; Score 225; DB 9; Length 1889;

```
Best Local Similarity 100.0%; Pred. No. 2.2e-64;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgacccagctgcccgcctatgcgaatgtgttgagccctctgg 60
Db 306 ATGCAGAGCTCTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 365
Qy 61 aggtattataaagcagagggcctatggagggccatgaggggctgaacgtcacagcaaca 120
Db 366 AGGATTATTAAGACGAGGGCCTATGGAGGCCCATGAGGGGCTGAAGCTCACACAACA 425
Qy 121 ggcgcagggcctgcccagcgcctttatttgcctgctacgaaaagttaaaaagacattg 180
Db 426 GCGCAGGGCCTGCCAGCCCTTTATTTGCTGCTACGAAAAGTTAAAAAGACATTG 485
Qy 181 agtcatgtaataccaccctgggggcaatagccatattgccaatggt 225
Db 486 AGTGATGTAATCCACCTGGGGCAATAGCCATATTGCCAATGGT 530

RESULT 6
HSA303078
ID HSA303078 standard: RNA; HUM; 1889 BP.
XX AC
XX AC
XX AC
SV AJ303078.1
XX AJ303078.1
XX 02-FEB-2001 (Rel. 66, Created)
DT 26-APR-2001 (Rel. 67, Last updated, Version 2)
XX
XX Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMRS3/4
DE gene), 1889 bp cDNA splice variant
DE
XX HMRS3/4 gene; mitochondrial RNA splicing protein 3/4.
XX
XX Homo sapiens (human)
OS
XX Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
XX [1]
RN 1-1889
RA Nikali K.;
RT ;
RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.
RL Nikali K., Human Molecular Genetics, National Public Health Institute,
RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.
XX
XX [3]
RN Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
RA Larsson C., Suomalainen A.;
RT "Characterization of a novel human putative mitochondrial transporter
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";
RL FEBS Lett. 494:79-84(2001).
XX
XX Key Location/Qualifiers
FH
FH 1. .1889
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT CDS 1032..1565
FT /gene="HMRS3/4"
FT /product="mitochondrial RNA splicing protein 3/4"
FT /function="putative mitochondrial solute carrier"
FT /protein_id="CAC27997.1"
FT /translation="MNPAEVVKRMQYNPSYHRVDCVRAVQWNEGAFYRSYITQL
FT TMNVPFOAHFMFYEFQAEHFNPNSSHLVSGACAGAAAATPDLVDCKTLNT
FT QESLALNSHTIGTITGMASAFRTYVQGGVTAYFRGVQARVIYQIPSTAIWMSYIEFFK
FT YLITKROEWRAGK"
XX
XX Sequence 1889 BP; 404 A; 508 C; 544 G; 433 T; 0 other;
```

```
Query Match 78.9%; Score 225; DB 17; Length 1889;
Best Local Similarity 100.0%; Pred. No. 2.2e-64;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgacccagcctgcccgcctatgcgaatgtgttgagccctctgg 60
Db 306 ATGCAGAGCTCTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 365
Qy 61 aggtattataaagcagagggcctatggagggccatgaggggctgaacgtcacagcaaca 120
Db 366 AGGATTATTAAGACGAGGGCCTATGGAGGCCCATGAGGGGCTGAAGCTCACACAACA 425
Qy 121 ggcgcagggcctgcccagcgcctttatttgcctgctacgaaaagttaaaaagacattg 180
Db 426 GCGCAGGGCCTGCCAGCCCTTTATTTGCTGCTACGAAAAGTTAAAAAGACATTG 485
Qy 181 agtcatgtaataccaccctgggggcaatagccatattgccaatggt 225
Db 486 AGTGATGTAATCCACCTGGGGCAATAGCCATATTGCCAATGGT 530

RESULT 7
AL353719/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete
sequence.
ACCESSION AL353719
VERSION AL353719.10
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Ramsay, H.
TITLE 1 (bases 1 to 123160)
JOURNAL Direct Submission
COMMENT Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 26, 2001 this sequence version replaced gi:14280413.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em; EMBL; Sw.;
SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-85A1 is from the library RPC1-11.1 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-85A1. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP11-85A1 is at 1 in this sequence. The
true left end of clone RP11-483F11 is at 123061 in this sequence.
The true right end of clone RP11-129J12 is at 51589 in this
sequence.
```

Location/Qualifiers	
1. .123160	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="10"	
/clone="RP11-85A1"	
/clone_lib="RPC1-11.1"	
complement(7066..7118)	
/note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."	
52162	
52167	
BASE COUNT	33221 a 28084 c 27196 g 34659 t
ORIGIN	
Query Match	
Best Local Similarity 100.0%; Pred. No. 4.2e-64;	
Matches 225;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1	atgcagagctacagctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
Db 85540	ATGCAGAGCTACAGCTGACCGTACCGCTGCGCTATCGCAATGTGTGAGGCCCTCTGG 85481
QY 61	aggattataaagacgagggccctatggagggcccatgaggggctgaacgtcacacca 120
Db 85480	AGGATTATAAGAACGAGGGCCCTATGGAGGCCCATGAGGGGCTGAAGCTCACGCAACA 85421
QY 121	ggcgcagggccctccacgcccctttatgttcctctcagaaagttaaaaaagacattg 180
Db 85420	GCGCAGGGCCCTGCGCCACGCGCCCTTTATTTCCTGCTAGGAAAGTTAAAAAGACATTG 85361
QY 181	agtgatgaatccaccctgggggccaatgccatattgccaatggt 225
Db 85360	AGTGATGTAATCCACCTGGGGGCATAGCCATATTGCCAATGGT 85316
RESULT 8	
AC096351/c	
LOCUS	AC096351 221062 bp DNA linear HTG 20-DEC-2001
DEFINITION	Rattus norvegicus chromosome Rf1 clone CH230-24M6, WORKING DRAFT
SEQUENCE	SEQUENCE, 33 unordered pieces..
AC096351	
AC096351.2	GI:17944054
VERSION	HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS	Norway rat.
SOURCE	Rattus norvegicus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 221062)
AUTHORS	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Albrooks, S.L., Amarantunge, H.C., Are, J.R., Banks, T., Barbara, J., Benton, J., Blinage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brivea, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,

Loulseged, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mahoney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Watlington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.	Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: GEZD Center clone name: CH230-24M6 ----- Summary Statistics Assembly program: Phrap; version 0.990329First call to findPhrapList Consensus quality: 200164 bases at least Q40 Consensus quality: 205303 bases at least Q30 Consensus quality: 209366 bases at least Q20 Estimated insert size: 209880; sum-of-contigs estimation Quality coverage: 0x in Q20 bases; agarose-fp estimation Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation ----- * NOTE: Estimated insert size may differ from sequence length * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html) * NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
2 (bases 1 to 221062)	1
Unpublished	32851: contig of 32851 bp in length
Worley, K. C.	32951: gap of unknown length
Direct Submission	32952
Submitted (17-SEP-2001)	48984: contig of 16033 bp in length
of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	49084: gap of unknown length
On Dec 20, 2001 this sequence version replaced gi:15627972.	63815: contig of 14731 bp in length
----- Genome Center	63816
Center: Baylor College of Medicine	63915: gap of unknown length
Center code: BCM	76105: gap of unknown length
Web site: http://www.hgsc.bcm.tmc.edu/	76106
Contact: hgsc-help@bcm.tmc.edu	88841: contig of 12736 bp in length
----- Project Information	88842
Center project name: GEZD	100591: contig of 11650 bp in length
Center clone name: CH230-24M6	100592
----- Summary Statistics	111702: contig of 11011 bp in length
Assembly program: Phrap; version 0.990329First call to findPhrapList	111703
Consensus quality: 200164 bases at least Q40	111802: gap of unknown length
Consensus quality: 205303 bases at least Q30	122664: contig of 10862 bp in length
Consensus quality: 209366 bases at least Q20	122665
Estimated insert size: 209880; sum-of-contigs estimation	122764: gap of unknown length
Quality coverage: 0x in Q20 bases; agarose-fp estimation	122765
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation	129711: contig of 6947 bp in length
-----	129712
* NOTE: Estimated insert size may differ from sequence length	
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)	
* NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	

Loulseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, N., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, P., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 32851: contig of 32851 bp in length
* 32852 32951: gap of unknown length
* 32952 48984: contig of 16033 bp in length
* 48985 49084: gap of unknown length
* 49085 63815: contig of 14731 bp in length
* 63816 63915: gap of unknown length
* 63916 76005: contig of 12090 bp in length
* 76006 76105: gap of unknown length
* 76106 88841: contig of 12736 bp in length
* 88842 88941: gap of unknown length
* 88942 100591: contig of 11650 bp in length
* 100592 100691: gap of unknown length
* 100692 111702: contig of 11011 bp in length
* 111703 111802: gap of unknown length
* 111803 122664: contig of 10862 bp in length
* 122665 122764: gap of unknown length
* 122765 129711: contig of 6947 bp in length
* 129712 129811: gap of unknown length

Qy 181 agtqatgtaataccaccctgggggcaataagcattatgccaatggt 225
|||||
Db 49921 AGTGATGTAATCCACCCGGGGCAATAGCCATATGCGCAATGGT 49877
|||||

RESULT 9
AF288621 4112 bp mRNA linear ROD 27-OCT-2001
LOCUS Mus musculus mitochondrial solute carrier-like protein mRNA,
DEFINITION complete cds.
ACCESSION AF288621
VERSION AF288621.1 GI:16506177
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4112)
AUTHORS Li,Q.-Z., Ruan,Q.-G., Eckenrode,S., Shi,J.-D., Cruz,P., Wang,C.-Y.
and She,J.-X.
TITLE A new gene which is highly expressed in NOD mice spleen
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4112)
AUTHORS Li,Q.-Z., Shi,J.-D., Cruz,P., Echenrode,S., Ruan,Q.-G. and
Wang,C.-Y.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2000) Pathology, University of Florida, 1600 SW
Archer Road, Rm. D6-15, Gainesville, FL 32610, USA
FEATURES
source
1..4112
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS
116..1132
/note="50E12"
/codon_start=1
/product="mitochondrial solute carrier-like protein"
/protein_id="AAL23859.1"
/db_xref="GI:16506178"
/translation="MELRRGSGVGNQAGRRMDGDCRDGGCGSKDAGSEYENLPTSAS
VSTHTAGAMAGILLEHSIMYPDVSQTRMQSLNDPKARYTSIYGALKRIMHTGEFWR
PLRGLNVMMGAGPAHAMYFACYENMARTLVDFSHGNSHLANGVAGSNATLLHDAV
MNPAAVVKQLQMYNSQSAFSCIRTVWTEGLGAFYSYTTTQLTMTNIPFQSITHTIT
YEFLOEQNPDRDYNPQSHISGGLAGALAAATPLDVKTLTLTNENMALSLANYS
GRLSGMANAFRTVQLNGLAGYFKGIQARVYIQMPSTAISWSVYEFKYILTKRQLEN
RTLY"
BASE COUNT 970 a 1024 c 970 g 1148 t
ORIGIN

Query Match 48.8%; Score 139.2; DB 10; Length 4112;
Best Local Similarity 68.6%; Pred. No. 1.8e-35;
Matches 192; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgaccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
|||||
Db 332 ATGCAGAGTGTGTAATCCAGATCCCAAGCCCGGTATACAAGCATCTATGCGCCCTCAAG 391
|||||

Qy 61 aggattataaagcaggaggcctatggagcccatgaggggctgaacgtcacagcaaca 120
|||||
Db 392 AGGATCATGTCACATGAAGGCTTCTGGAGGCCCTCTGCGGGCCCTGAACGTGATGATG 451
|||||

Qy 121 ggcagaggcctgcccagcgccttattttgcctgctacgaaagttataaaagacattg 180
|||||
Db 452 GTCAGGGCCCGCGCAGCCATGTATTTTCCTGCTATGAAACATGAAAGGACTTTA 511
|||||

Qy 181 agtqatgtaataccaccctgggggcaataagcattatgccaatgtgttgaggccgtgtgtg 240
|||||
Db 512 AATGACGTTTTTCAGCCCAAGCAAGCAAGCCATCTAGCTAATGGGTAGCTGGGAGTATG 571
|||||

Qy 241 gcaacattactcatgatgacagccatgaacccctgcggaag 280
|||||
Db 572 GCCACCTACTCCAGATGTCAGAGTAATGAATCCAGCAGAAG 611
|||||

* 129812 137247: contig of 7436 bp in length
* 137248 137347: gap of unknown length
* 137348 144676: contig of 7329 bp in length
* 144677 144776: gap of unknown length
* 144777 152124: contig of 7348 bp in length
* 152125 152224: gap of unknown length
* 152225 158878: contig of 6654 bp in length
* 158879 158978: gap of unknown length
* 158979 163928: contig of 4950 bp in length
* 163929 164028: gap of unknown length
* 164029 169353: contig of 5325 bp in length
* 169354 169453: gap of unknown length
* 169454 175590: contig of 6137 bp in length
* 175591 175690: gap of unknown length
* 175691 180475: contig of 4785 bp in length
* 180476 180575: gap of unknown length
* 180576 184657: contig of 4082 bp in length
* 184658 184757: gap of unknown length
* 184758 190332: contig of 5575 bp in length
* 190333 190432: gap of unknown length
* 190433 194134: contig of 3702 bp in length
* 194135 194234: gap of unknown length
* 194235 196592: contig of 2358 bp in length
* 196593 196692: gap of unknown length
* 196693 200530: contig of 3838 bp in length
* 200531 200630: gap of unknown length
* 200631 202809: contig of 2179 bp in length
* 202810 202909: gap of unknown length
* 202910 205034: contig of 2125 bp in length
* 205035 205134: gap of unknown length
* 205135 207176: contig of 2042 bp in length
* 207177 207276: gap of unknown length
* 207277 208845: contig of 1569 bp in length
* 208846 208945: gap of unknown length
* 208946 211271: contig of 2326 bp in length
* 211272 211371: gap of unknown length
* 211372 213302: contig of 1931 bp in length
* 213303 213402: gap of unknown length
* 213403 215484: contig of 2082 bp in length
* 215485 215584: gap of unknown length
* 215585 215699: contig of 1385 bp in length
* 215699 216970: gap of unknown length
* 216970 217079: contig of 1289 bp in length
* 217079 218359: gap of unknown length
* 218359 218458: gap of unknown length
* 218459 219886: contig of 1428 bp in length
* 219887 219886: gap of unknown length
* 219887 221062: contig of 1076 bp in length.
* 219987

FEATURES
source
1..221062
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/chromosome="Rf1"
/clone="CH230-24M6"
BASE COUNT 58389 a 50577 c 49732 g 58877 t 3487 others
ORIGIN

Query Match 71.1%; Score 202.6; DB 2; Length 221062;
Best Local Similarity 93.8%; Pred. No. 1.6e-56;
Matches 211; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgaccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
|||||
Db 50101 ATGCAGAGTCTACAGCTGATGATCCAGCCCGCCCTATCGAACAGCTGTGGAGGCTCTCTGG 50042
|||||

Qy 61 aggattataaagcaggaggcctatggaggcccatgaggggctgaacgtcacagcaaca 120
|||||
Db 50041 AGAATTAAGAACAGAGGCGCTGTGGAGGCCCATGAGGGGCTGAACGTCACTGCGACT 49982
|||||

Qy 121 ggcagaggcctgcccagcgccttattttgcctgctacgaaagttataaaagacattg 180
|||||
Db 49981 GCGCGGGGCGCTGCCACGCGCCCTCTATTTTTCCTGCTACGAAAAAGTTAAAAAGACATTG 49922
|||||

RESULT	10	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	AUTHORS	TITLE	JOURNAL	COMMENT
AF223466	1	1429 bp	linear	PRI 17-APR-2000	AC099368	GI:17973175	HTG: HTGS_PHASE1	Norway rat.	Rattus norvegicus	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	1 (bases 1 to 163205)	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J., Benton, J., Bimage, K., Blankenbush, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyne, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flegg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, N., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogul, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.	Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: GHJ Center clone name: CH230-84G11 ----- Summary Statistics Assembly program: Phrap; version 0.990329First call to findPhrapList			
AF223466	2	1429 bp	linear	PRI 17-APR-2000	AC099368	GI:17973175	HTG: HTGS_PHASE1	Norway rat.	Rattus norvegicus	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	1 (bases 1 to 163205)	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J., Benton, J., Bimage, K., Blankenbush, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyne, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flegg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, N., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogul, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.	Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: GHJ Center clone name: CH230-84G11 ----- Summary Statistics Assembly program: Phrap; version 0.990329First call to findPhrapList			
AF223466	1	1429 bp	linear	PRI 17-APR-2000	AC099368	GI:17973175	HTG: HTGS_PHASE1	Norway rat.	Rattus norvegicus	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	1 (bases 1 to 163205)	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J., Benton, J., Bimage, K., Blankenbush, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyne, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flegg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, N., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogul, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson,				

Db	13898	GATGTAATCCACCCGGGGCCAAATAGCCATATTGCGCAATGGT	13939	
RESULT	12			
AF361699				
LOCUS				
DEFINITION		5869 bp mRNA linear ROD 07-NOV-2001		
ACCESSION	AF361699	Mus musculus mitochondrial carrier-like protein mRNA, complete cds;		
VERSION	AF361699	nuclear gene for mitochondrial product.		
KEYWORDS	AF361699.1	GI:16755527		
SOURCE		house mouse.		
ORGANISM		Mus musculus		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
1	(bases 1 to 5869)			
Li.O., Eckenrode.S., Wang.C., Ruan.Q., Shi.J., McIndoe.R.A. and				
She.J.				
TITLE		A novel mouse mitochondrial carrier protein gene is up-regulated		
JOURNAL		from young to adult NOD mice		
REFERENCE		Unpublished		
AUTHORS		2 (bases 1 to 5869)		
Li.O., Eckenrode.S., Wang.C., Ruan.Q., Shi.J., McIndoe.R.A. and				
She.J.				
TITLE		Direct Submission		
JOURNAL		Submitted (15-MAR-2001) Department of Pathology, Immunology and		
1		Laboratory Sciences, University of Florida, 1600 SW Archer Road,		
Room D6-15, Gainesville, FL 32610, USA				
Location/Qualifiers				
FEATURES				
source		1..5869		
		/organism="Mus musculus"		
		/strain="C57BL/6J"		
		/db_xref="taxon:10090"		
		/chromosome="14"		
		/map="between D14Mit124 and D14Mit236"		
		116..684		
		/note="MCLP"		
		/codon_start=1		
		/evidence=not_experimental		
		/product="mitochondrial carrier-like protein"		
		/protein_id="AAL27990.1"		
		/db_xref="GI:16755528"		
		/translation="MELRRGVGNQAGRRMDGDRDGGCGSKDAGSDYENLPFSAS		
		VSTHMTAGAMGILESHYPVDSYKTRQSLNPDPKARYTYSIYGALKRIMHTEGFWR		
		PLRLGNVMMGAGPAHAMYFACYENKRTLNDFVSHQGNHLANGILKAFVMSWEALL		
		SGASSPGFSLNLPROTENSRTV"		
BASE COUNT	1387 a	1427 c	1390 g	1665 t
ORIGIN				
Query Match		38.5%;	Score 109.8;	DB 10; Length 5869;
Best Local Similarity		68.0%;	Pred. No. 1.5e-25;	
Matches 153;	Conservative	0;	Mismatches 72;	Indels 0; Gaps 0;
QY	1	atgcagagtctacagctgaccagctgcccagctgcgcgtatcgcaatgtgtgaggccctctgg	60	
Db	332	ATGCAGAGTTTGAATCCAGATCCAAAGCCGGTATACAAAGCATCTATGCGGCCCTCAAG	391	
QY	61	agattataaagcagcagcctatgagcccatgaggggagctgaacctcacagcaaca	120	
Db	392	AGATCATGCACACTGAAGGCTTCTGGAGGCCCCCTCGGGGGCTGAACGTGATGATG	451	
QY	121	ggcgcaggccctgccacagccctttatttgctgctacgaaagttaaaaaagacattg	180	
Db	452	GGTGACGGCCCGCGCACGCCATGATTTTGGCTGCTATGAACATGAAGAGGACTTTTA	511	
QY	181	agtgtgttaatccacccctggggcaatagcccatattggcaatgggt	225	
Db	512	AATGACGTTTTTCAGCCACCAAGAAACAGACCCATCTAGCTAATGGT	556	
RESULT	13			
AC108878				

LOCUS	AC108878	191702 bp	DNA	linear	HTC 06-FEB-2002
DEFINITION		Mus musculus clone RP23-256H24 strain C57BL6/J, *** SEQUENCING IN			
PROGRESS	***	17 unordered pieces.			
ACCESSION	AC108878				
VERSION	AC108878.2	GI:18542954			
KEYWORDS	HTG; HTGS-PHASE1.				
SOURCE		house mouse.			
ORGANISM		Mus musculus			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1	(bases 1 to 191702)				
Grills.G., Li.L., Montgomery.K.T., Brown.W.A., Chiu.D., Decker.J.,					
Fusina.M., Haider.A., Keller.A., Perera.A., Shim.C., Thomas.E.,					
Zencheck.W., Xi.C., Juels.P. and Kucherlapati.R.					
TITLE		High Throughput Mouse Sequencing			
JOURNAL		Unpublished			
REFERENCE		2 (bases 1 to 191702)			
AUTHORS		Grills.G., Li.L., Montgomery.K.T., Brown.W.A., Chiu.D., Decker.J.,			
1		Fusina.M., Haider.A., Keller.A., Perera.A., Shim.C., Thomas.E.,			
Zencheck.W., Xi.C., Juels.P. and Kucherlapati.R.					
TITLE		Direct Submission			
JOURNAL		Submitted (01-FEB-2002) Harvard Partners Center for Genetics and			
1		Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA			
02139, USA					
COMMENT		On Feb 6, 2002 this sequence version replaced gi:18464039.			
		-----Genome Center			
		Center: Harvard Partners Genome Center			
		Center Code: HPGC			
		Web site:			
		http://wchanning.bwh.harvard.edu:9088/hpcgg/jsp/hpcgg/Sequence/mous			
		e.html			
		Contact: gntm@capecod.bwh.harvard.edu			
		-----Summary Statistics			
		Center project name: AKY			
		Sequencing vector: pUC18; L08752			
		Chemistry: Dye-terminator Big Dye; 100%			
		*Consensus quality: 187211 at least Q20			
		*Consensus quality: 185727 at least Q30			
		*Consensus quality: 183181 at least Q40			
		Estimated insert size: agarose-PP - N/A			
		**Estimated insert size: 191382 - sum-of-contigs			
		Quality coverage: agarose-PP - N/A			
		Quality coverage: 5.6 x in Q20 bases; sum-of-contigs estimation			

		* NOTE: This is a 'working draft' sequence. It currently			
		* consists of 17 contigs. The true order of the pieces			
		* is not known and their order in this sequence record is			
		* arbitrary. Gaps between the contigs are represented as			
		* runs of N, but the exact sizes of the gaps are unknown.			
		* This record will be updated with the finished sequence			
		* as soon as it is available and the accession number will			
		* be preserved.			
		* 1 48676: contig of 48676 bp in length			
		* 48677 48696: gap of unknown length			
		* 48697 89065: contig of 40369 bp in length			
		* 89066 89085: gap of unknown length			
		* 89086 104994: contig of 15909 bp in length			
		* 104995 105014: gap of unknown length			
		* 105015 119612: contig of 14598 bp in length			
		* 119613 119632: gap of unknown length			
		* 119633 135969: contig of 16337 bp in length			
		* 135970 135989: gap of unknown length			
		* 135990 151803: contig of 15814 bp in length			
		* 151804 151823: gap of unknown length			
		* 151824 160052: contig of 8229 bp in length			
		* 160053 160072: gap of unknown length			
		* 160073 169021: contig of 8949 bp in length			
		* 169022 169041: gap of unknown length			
		* 169042 175745: contig of 6704 bp in length			
		* 175746 175765: gap of unknown length			
		* 175766 180321: contig of 4536 bp in length			
		* 180322 180341: gap of unknown length			
		* 180342 182588: contig of 2247 bp in length			
		* 182589 182588: contig of 2247 bp in length			


```
* 182589 182608: gap of unknown length
* 182609 183975: contig of 1367 bp in length
* 183976 183995: gap of unknown length
* 183996 186393: contig of 2398 bp in length
* 186394 186413: gap of unknown length
* 186414 188416: contig of 2003 bp in length
* 188417 188436: gap of unknown length
* 188437 189584: contig of 1148 bp in length
* 189585 189604: gap of unknown length
* 189605 190648: contig of 1044 bp in length
* 190649 190668: gap of unknown length
* 190669 191702: contig of 1034 bp in length.
FEATURES
    Location/Qualifiers
        1..191702
            /organism="Mus musculus"
            /strain="C57BL6/J"
            /db_xref="taxon:10090"
            /clone="RP23-256H24"
            /sex="male"
BASE COUNT 53763 a 43806 c 42665 g 51108 t 360 others
ORIGIN
```

```
1 atgcagagtctacacgctgaccacgctcccgctatcgaatgtgtggagccctctgg 60
|||||
Db 106283 ATGCAGAGTTTAAATCCAGATCCCAAGCCGGTATACAGCATATATGGCGCCCTCAAG 106342
|||||
* 61 aggattataagaacgagggcctatggagggccatgagggcgtgaacgtcacagcaaca 120
|||||
Db 106343 AGGATCATGCACACTGAGGCTTCTGGAGGCCCTCGGGGCTGAAGCTGATGATG 106402
|||||
Qy 121 ggcgcagggcctgcacacgccccttatttgcctgctctacgaaagttaaagaagacattg 180
|||||
Db 106403 GGTGCGAGGCGCCGCGCAGCCCATGTATTGTGCTGCTATGAAGACATGAAAGGACTTTA 106462
|||||
Qy 181 agtgaatgaatccaccctgggggcaatagccatatgtccaatggt 225
|||||
Db 106463 AATGACGTTTTTCAGCCACCAAGAAACAGCCCATCTAGCTAATGTT 106507
|||||
```

```
RESULT 14
AX306258
LOCUS AX306258 483 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 1009 from Patent WO0188188.
ACCESSION AX306258
VERSION AX306258.1 GI:17645513
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (sites)
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 1009 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
    Location/Qualifiers
        1..483
            /organism="Mus musculus"
            /db_xref="taxon:10090"
BASE COUNT 127 a 139 c 121 g 96 t
ORIGIN
```

```
Query Match 36.9%; Score 105.2; DB 6; Length 483;
Best Local Similarity 93.2%; Pred. No. 3.5e-24;
Matches 110; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 165 gttaaaaaagacattgadtgatgtatccaccctgggggcaatagccattgccaattg 224
|||||
```

```
Db 15 GTTAAAAAGACATTGAGTGACGTAATCACCCAGGGCGCAATAGCCATATTGCAATGG 74
|||||
Qy 225 tcgcgccgggtgtggcaacattactctcatgataccagcatgaacctgggaagcc 282
|||||
Db 75 TGCAGCCGGATGTGTGGCGACATTACTTTCATGATGACGCCATGAATCCAGCGGAAGTC 132
|||||
```

```
RESULT 15
AF216674
LOCUS AF216674 108765 bp DNA linear HTG 02-JUN-2001
DEFINITION Homo sapiens chromosome 8 clone RP1-158P13 map 8p, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
ACCESSION AF216674
VERSION AF216674.3 GI:14280182
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 108765)
AUTHORS Schilhabel,M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Siddiqui,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,
Rosenthal,A. and Platzer,M.
TITLE Chromosome 8 genomic sequence
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 108765)
AUTHORS Polley,A., Wen,G., Baumgart,C., Dette,M., Jahn,N., Schilhabel,M.,
Menzel,U. and Rosenthal,A.
JOURNAL Direct Submission
COMMENT Submitted (15-DEC-1999) Genome Analysis, Institute of Molecular
Biotechnology, Reutenbergstrasse 11, Jena 07745, Germany
On Jun 2, 2001 this sequence version replaced gi:8151796.
-----
Center: Insitute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
-----
Project Information
Center project name: H241
Center clone name: RP1-158P13
-----
Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 87772 bases at least Q40
Consensus quality: 95159 bases at least Q30
Consensus quality: 100612 bases at least Q20
Quality coverage: 3.95 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1101: contig of 1101 bp in length
* 1102 1201: gap of unknown length
* 1202 21139: contig of 19938 bp in length
* 21140 21239: gap of unknown length
* 21240 36201: contig of 14962 bp in length
* 36202 36301: gap of unknown length
* 36302 47101: contig of 10800 bp in length
* 47102 47201: gap of unknown length
* 47202 56806: contig of 9605 bp in length
* 56807 56906: gap of unknown length
* 56907 65083: contig of 8177 bp in length
* 65084 65183: gap of unknown length
* 65184 73621: contig of 8338 bp in length
* 73622 73522: gap of unknown length
```


THIS PAGE BLANK (USPTO)


```
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
PA
XX Shlmkets RA, Leach M;
XX
XX WPT: 2000-602362/57.
DR P-PSDB; AAB40634.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 5; Page 853-854; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulneryary;
CC antiporiatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus, lupus
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1418 BP; 353 A; 400 C; 368 G; 297 T; 0 other;
SQ

Query Match 100.0%; Score 285; DB 21; Length 1418;
Best Local Similarity 100.0%; Pred. No. 9.3e-87;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgacccagctgccgctatcgcaatgtgttgaggccctctgg 60
Db 178 atgcagagtctacagcctgacccagctgccgctatcgcaatgtgttgaggccctctgg 237
Qy 61 aggattataagaacgaggccctatgaggcccatgaggggctgaacgtcacagcaaca 120
Db 238 aggattataagaacgaggccctatgaggcccatgaggggctgaacgtcacagcaaca 297
Qy 121 ggcgcaggcctgccacgcctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 298 ggcgcaggcctgccacgcctttatttgcctgctacgaaagttaaaaaagacattg 357
Qy 181 agtgatgaatccacctgggggcaatagccatttcccaatggtgcggcggtgtgtg 240
Db 358 agtgatgaatccacctgggggcaatagccatttcccaatggtgcggcggtgtgtg 417
Qy 241 gcaacattactctatgatgcagccatgaacccctgcgggaagctga 285
Db 418 gcaacattactctatgatgcagccatgaacccctgcgggaagctga 462

RESULT 2
AAI60661/c
ID AAI60661 standard; cDNA; 1294 BP.
XX
```

```
AC AAI60661;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 4650.
DE
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-052317.
PR
XX 09-JUL-2000; 2000US-0598042.
PR
XX 19-JUL-2000; 2000US-0620312.
PR
XX 14-AUG-2000; 2000US-0653450.
PR
XX 14-SEP-2000; 2000US-0662191.
PR
XX 19-OCT-2000; 2000US-0693036.
PR
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR
XX P-PSDB; AAM41505.
DR
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 4650; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 1294 BP; 267 A; 335 C; 368 G; 324 T; 0 other;
SQ

Query Match 98.2%; Score 280; DB 22; Length 1294;
Best Local Similarity 100.0%; Pred. No. 4.5e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgacccagctgcccgctatcgcaatgtgttgaggccctctgg 60
Db 1134 ATGCAGAGTCTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 1075
Qy 61 aggattataagaacgaggccctatgaggcccatgaggggctgaacgtcacagcaaca 120
```



```
Query Match      98.2%; Score 280; DB 21; Length 1336;
Best Local Similarity 100.0%; Pred. No. 4.6e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atcgagagctacagcctgacccagctccgcgtatcgcaatgtgttgaggccctctgg 60
Db 197 atcgagagctacagcctgacccagctccgcgtatcgcaatgtgttgaggccctctgg 256

QY 61 aggattataagaagaggcctatggagggccatgaggggctgaacgtcacagcaaca 120
Db 257 aggattataagaagaggcctatggagggccatgaggggctgaacgtcacagcaaca 316

QY 121 ggcgcaggcctgcccacgcctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 317 ggcgcaggcctgcccacgcctttatttgcctgctacgaaagttaaaaaagacattg 376

QY 181 agtgaatgaatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtg 240
Db 377 agtgaatgaatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtg 436

QY 241 gcaacattacttcattgatgcagccatgaacctgcggaag 280
Db 437 gcaacattacttcattgatgcagccatgaacctgcggaag 476

RESULT 5
AAF66156
ID AAF66156 standard; cDNA; 401 BP.
XX
AC AAF66156;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1912.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
XX
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 818; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
```

```
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match      95.8%; Score 273; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 6.5e-83;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 gtctacagcctgacccagctccgcgtatcgcaatgtgttgaggccctctggaggatta 67
Db 9 gtctacagcctgacccagctccgcgtatcgcaatgtgttgaggccctctggaggatta 68

QY 68 taagaacggaggcctatggagggccatgaggggctgaacgtcacagcaacagcgccag 127
Db 69 taagaacggaggcctatggagggccatgaggggctgaacgtcacagcaacagcgccag 128

QY 128 ggcctgcccacgcctttatttgcctgctacgaaaagttaaaaaagacattgagtgtg 187
Db 129 ggcctgcccacgcctttatttgcctgctacgaaaagttaaaaaagacattgagtgtg 188

QY 188 taatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtgccaacat 247
Db 189 taatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtgccaacat 248

QY 248 tacttcattgatgcagccatgaacctgcggaag 280
Db 249 tacttcattgatgcagccatgaacctgcggaag 281

RESULT 6
AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX
AC AAI58875;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1078.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
XX
PR 25-APR-2000; 2000US-0552317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.
XX
PR 03-AUG-2000; 2000US-0653450.
XX
PR 14-SEP-2000; 2000US-0662191.
XX
PR 19-OCT-2000; 2000US-0693036.
XX
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
```

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
DR WPI: 2001-442253/47.
DR P-PSDB; AAM39719.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Claim 1; SEQ ID NO 1078; 10078pp; English.
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX SQ Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;
Query Match 95.4%; Score 272; DB 22; Length 1316;
Best Local Similarity 98.2%; Pred. No. 2,4e-82;
Matches 275; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 atgcagagctacagcctgaccagctgccgctatcgcaatgtgttgaggccctctgg 60
Db 176 atgcagagctacagctgaccagctgcccagctgcccagctgcaatgtgttgaggctctctgg 235
Qy 61 aggtattataaagcagggccctatgaggccatgaggggctgaacgtcacacaaca 120
Db 236 aggtattataaagcagggccctatgaggccatgaggggctgaacgtcacacaaca 295
Qy 121 ggcgcaggcctgccacgcctttatttgcctctacgaaaagttaaaaaagacattg 180
Db 296 ggcgcaggcctgccacgcctttatttgcctctacgaaaagttaaaaaagacattg 355
Qy 181 agtgatgtaataccacctgggggcaatagccaattgccaatggtgcgcccgggtgtgtg 240
Db 356 agtgatgtaataccacctgggggcaatagccaattgccaatggtgcgcccgggtgtgtg 415
Qy 241 gcaacattacttcattgatgcagccatgaacctgcggaag 280
Db 416 gcaacattacttcattgatgcagccatgaacctgcggaag 455
RESULT 7
AAS03906
ID AAS03906 standard; cDNA; 2502 BP.
XX AA AAS03906;
XX Human secreted protein gene #25.
DT 29-AUG-2001 (first entry)
DE Human secreted protein gene #25.
KW Human secreted protein; autoimmune disorder; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;
KW nervous system disorder; bacterial infection; viral infection; ss;
KW fungal infection; ocular disorder; wound healing; tissue regeneration;
KW epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.
XX

OS Homo sapiens.
XX WO200123598-A1.
PN 05-APR-2001.
XX 26-SEP-2000; 2000WO-US26324.
PF 27-SEP-1999; 99US-0155807.
PR (HUMA-) HUMAN GENOME SCI INC.
PA Komatsoulis G, Ruben SM, Rosen CA;
XX WPI: 2001-281684/29.
DR P-PSDB; AAU01950, AAU01989.
XX Forty one nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX Disclosure; Page 460-461; 518pp; English.
XX Sequences AAS03873-AAS03922 represent isolated nucleic acid molecules and
CC PCR primers of the invention. acid of the invention. Secreted proteins
CC and their related nucleic acids can be used in the diagnosis of or
CC susceptibility to a pathological condition by determining the presence or
CC absence of a mutation in a nucleic acid or the presence or amount of
CC expression of a secreted protein. The sequences are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. The antibodies to the
CC polypeptides can also be used in alleviating symptoms associated with
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). The disorders include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The peptides can also be used to aid wound healing and epithelial cell
CC proliferation, to help prevent skin ageing due to sunburn, to maintain
CC organs before transplantation, to regenerate tissues, in chemotaxis and
CC as a food additive or preservative to alter storage capabilities.
XX SQ Sequence 2502 BP; 587 A; 655 C; 669 G; 589 T; 2 other;
Query Match 78.9%; Score 225; DB 22; Length 2502;
Best Local Similarity 100.0%; Pred. No. 3,4e-66;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 atgcagagctacagcctgaccagctgccgctatcgcaatgtgttgaggccctctgg 60
Db 195 atgcagagctacagcctgaccagctgccgctatcgcaatgtgttgaggccctctgg 254
Qy 61 aggtattataaagcagggccctatgaggccatgaggggctgaacgtcacacaaca 120
Db 255 aggtattataaagcagggccctatgaggccatgaggggctgaacgtcacacaaca 314
Qy 121 ggcgcaggcctgccacgcctttatttgcctctacgaaaagttaaaaaagacattg 180
Db 315 ggcgcaggcctgccacgcctttatttgcctctacgaaaagttaaaaaagacattg 374
Qy 181 agtgatgtaataccacctgggggcaatagccaattgccaatggt 225
Db 375 agtgatgtaataccacctgggggcaatagccaattgccaatggt 419
RESULT 8
AAK69779
ID AAK69779 standard; DNA; 27960 BP.
XX AA AAK69779;
AC

XX	06-NOV-2001	(first entry)
DT	Human immune/haematopoietic antigen genomic sequence	SEQ ID NO:24591.
XX	Human immune/haematopoietic antigen genomic sequence	SEQ ID NO:24591.
DE	Human immune/haematopoietic antigen genomic sequence	SEQ ID NO:24591.
XX	Human immune/haematopoietic antigen genomic sequence	SEQ ID NO:24591.
XX	Human immune/haematopoietic antigen genomic sequence	SEQ ID NO:24591.
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;	
KW	cystatic; gene therapy; vaccine; metastasis; ds.	
XX	Homo sapiens.	
OS	WO200157182-A2.	
XX	WO200157182-A2.	
PN	09-AUG-2001.	
XX	09-AUG-2001.	
PD	17-JAN-2001; 2001WO-US01354.	
XX	17-JAN-2001; 2001WO-US01354.	
XX	31-JAN-2000; 2000US-0179065.	
PR	31-JAN-2000; 2000US-0179065.	
PR	04-FEB-2000; 2000US-0180628.	
PR	04-FEB-2000; 2000US-0180628.	
PR	24-FEB-2000; 2000US-0184664.	
PR	24-FEB-2000; 2000US-0184664.	
PR	02-MAR-2000; 2000US-0186350.	
PR	02-MAR-2000; 2000US-0186350.	
PR	16-MAR-2000; 2000US-0189874.	
PR	16-MAR-2000; 2000US-0189874.	
PR	17-MAR-2000; 2000US-0190076.	
PR	17-MAR-2000; 2000US-0190076.	
PR	18-APR-2000; 2000US-0198123.	
PR	18-APR-2000; 2000US-0198123.	
PR	19-MAY-2000; 2000US-0205515.	
PR	19-MAY-2000; 2000US-0205515.	
PR	07-JUN-2000; 2000US-0209467.	
PR	07-JUN-2000; 2000US-0209467.	
PR	28-JUN-2000; 2000US-0214886.	
PR	28-JUN-2000; 2000US-0214886.	
PR	30-JUN-2000; 2000US-0215135.	
PR	30-JUN-2000; 2000US-0215135.	
PR	07-JUL-2000; 2000US-0216647.	
PR	07-JUL-2000; 2000US-0216647.	
PR	07-JUL-2000; 2000US-0216880.	
PR	07-JUL-2000; 2000US-0216880.	
PR	11-JUL-2000; 2000US-0217487.	
PR	11-JUL-2000; 2000US-0217487.	
PR	11-JUL-2000; 2000US-0217496.	
PR	11-JUL-2000; 2000US-0217496.	
PR	14-JUL-2000; 2000US-0218290.	
PR	14-JUL-2000; 2000US-0218290.	
PR	26-JUL-2000; 2000US-0220963.	
PR	26-JUL-2000; 2000US-0220963.	
PR	26-JUL-2000; 2000US-0220964.	
PR	26-JUL-2000; 2000US-0220964.	
PR	14-AUG-2000; 2000US-0224518.	
PR	14-AUG-2000; 2000US-0224518.	
PR	14-AUG-2000; 2000US-0224519.	
PR	14-AUG-2000; 2000US-0224519.	
PR	14-AUG-2000; 2000US-0225213.	
PR	14-AUG-2000; 2000US-0225213.	
PR	14-AUG-2000; 2000US-0225214.	
PR	14-AUG-2000; 2000US-0225214.	
PR	14-AUG-2000; 2000US-0225266.	
PR	14-AUG-2000; 2000US-0225266.	
PR	14-AUG-2000; 2000US-0225267.	
PR	14-AUG-2000; 2000US-0225267.	
PR	14-AUG-2000; 2000US-0225268.	
PR	14-AUG-2000; 2000US-0225268.	
PR	14-AUG-2000; 2000US-0225270.	
PR	14-AUG-2000; 2000US-0225270.	
PR	14-AUG-2000; 2000US-0225447.	
PR	14-AUG-2000; 2000US-0225447.	
PR	14-AUG-2000; 2000US-0225757.	
PR	14-AUG-2000; 2000US-0225757.	
PR	14-AUG-2000; 2000US-0225758.	
PR	14-AUG-2000; 2000US-0225758.	
PR	14-AUG-2000; 2000US-0225759.	
PR	14-AUG-2000; 2000US-0225759.	
PR	18-AUG-2000; 2000US-0226279.	
PR	18-AUG-2000; 2000US-0226279.	
PR	22-AUG-2000; 2000US-0226681.	
PR	22-AUG-2000; 2000US-0226681.	
PR	22-AUG-2000; 2000US-0226868.	
PR	22-AUG-2000; 2000US-0226868.	
PR	22-AUG-2000; 2000US-0227182.	
PR	22-AUG-2000; 2000US-0227182.	
PR	23-AUG-2000; 2000US-0227009.	
PR	23-AUG-2000; 2000US-0227009.	
PR	30-AUG-2000; 2000US-0228924.	
PR	30-AUG-2000; 2000US-0228924.	
PR	01-SEP-2000; 2000US-0229287.	
PR	01-SEP-2000; 2000US-0229287.	
PR	01-SEP-2000; 2000US-0229343.	
PR	01-SEP-2000; 2000US-0229343.	
PR	01-SEP-2000; 2000US-0229345.	
PR	01-SEP-2000; 2000US-0229345.	
PR	05-SEP-2000; 2000US-0229509.	
PR	05-SEP-2000; 2000US-0229509.	
PR	05-SEP-2000; 2000US-0229513.	
PR	05-SEP-2000; 2000US-0229513.	
PR	06-SEP-2000; 2000US-0230437.	

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis.
XX
XX
PS Disclosure; SEQ ID NO 24591; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;
SQ

Query Match 78.9%; Score 225; DB 22; Length 27960;
Best Local Similarity 100.0%; Pred. No. 9.8e-66;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagtctacagctgaccagctgcccgtatcgcaatgtgtggagccctctgg 60
|||||
Db 18295 atgcagagtctacagctgaccagctgcccgtatcgcaatgtgtggagccctctgg 18354
|||||

QY 61 aggattataagaacgaggggccctatggagcccatgagggggctgaacgtcacagcaaca 120
|||||
Db 18355 aggattataagaacgaggggccctatggagcccatgagggggctgaacgtcacagcaaca 18414
|||||

QY 121 ggcgcaggccctgccacgcccctttatttgcctgtctagaaagttaaaaaagacattg 180
|||||
Db 18415 ggcgcaggccctgccacgcccctttatttgcctgtctagaaagttaaaaaagacattg 18474
|||||

QY 181 agtgcgttaataccacctgggggcaatgccatattgccaatgg 225
|||||
Db 18475 agtgcgttaataccacctgggggcaatgccatattgccaatgg 18519
|||||

RESULT 9
AAK73320
ID AAK73320 standard; DNA; 27960 BP.
XX
AC AAK73320;
XX
XX 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28132.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
KW

XX Homo sapiens.
OS WO200157182-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01354.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0228868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.

```

PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241617.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259676.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 28132; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;

Query Match 78.9%; Score 225; DB 22; Length 27960;
Best Local Similarity 100.0%; Pred. No. 9.8e-66;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctctacagctgacccagctgcccgcctatgagcccatgagggcgtgaacgtcacagcaaca 120
Db 18295 atgcagagctctacagctgacccagctgcccgcctatgagcccatgagggcgtgaacgtcacagcaaca 18354
QY 61 aggattataagaacgagggcctatgagcccatgagggcgtgaacgtcacagcaaca 120
Db 18355 aggattataagaacgagggcctatgagcccatgagggcgtgaacgtcacagcaaca 18414
QY 121 ggcgcagggcctgcccacgcccctttatttgcctgctacgaaagttaaaaagacattg 180
Db 18415 ggcgcagggcctgcccacgcccctttatttgcctgctacgaaagttaaaaagacattg 18474
QY 181 agtgatgtaataccaccctgggggccaatagcccatattgccaatggt 225
Db 18475 agtgatgtaataccaccctgggggccaatagcccatattgccaatggt 18519

RESULT 10
AAK58672
ID AAK58672 standard; cDNA; 339 BP.
XX
XX AAK58672;
AC AAK58672;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/hematopoietic antigen encoding cDNA SEQ ID NO:3732.
XX
KW Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
```

PF	17-JAN-2001; 2001WO-US01354.	PR	29-SEP-2000; 2000US-0236370.
XX		PR	02-OCT-2000; 2000US-0236802.
PR	31-JAN-2000; 2000US-0179065.	PR	02-OCT-2000; 2000US-0237037.
PR	04-FEB-2000; 2000US-0180628.	PR	02-OCT-2000; 2000US-0237038.
PR	24-FEB-2000; 2000US-0184664.	PR	02-OCT-2000; 2000US-0237039.
PR	02-MAR-2000; 2000US-0186350.	PR	02-OCT-2000; 2000US-0237040.
PR	16-MAR-2000; 2000US-0189874.	PR	13-OCT-2000; 2000US-0239935.
PR	17-MAR-2000; 2000US-0190076.	PR	13-OCT-2000; 2000US-0239937.
PR	18-APR-2000; 2000US-0198123.	PR	20-OCT-2000; 2000US-0240960.
PR	19-MAY-2000; 2000US-0205515.	PR	20-OCT-2000; 2000US-0241221.
PR	07-JUN-2000; 2000US-0209467.	PR	20-OCT-2000; 2000US-0241785.
PR	28-JUN-2000; 2000US-0214886.	PR	20-OCT-2000; 2000US-0241786.
PR	30-JUN-2000; 2000US-0215135.	PR	20-OCT-2000; 2000US-0241787.
PR	07-JUL-2000; 2000US-0215647.	PR	20-OCT-2000; 2000US-0241808.
PR	07-JUL-2000; 2000US-0216880.	PR	20-OCT-2000; 2000US-0241809.
PR	11-JUL-2000; 2000US-0217487.	PR	20-OCT-2000; 2000US-0241826.
PR	11-JUL-2000; 2000US-0217496.	PR	01-NOV-2000; 2000US-0244617.
PR	14-JUL-2000; 2000US-0218290.	PR	08-NOV-2000; 2000US-0246474.
PR	26-JUL-2000; 2000US-0220963.	PR	08-NOV-2000; 2000US-0246475.
PR	26-JUL-2000; 2000US-0220964.	PR	08-NOV-2000; 2000US-0246476.
PR	14-AUG-2000; 2000US-0224518.	PR	08-NOV-2000; 2000US-0246477.
PR	14-AUG-2000; 2000US-0224519.	PR	08-NOV-2000; 2000US-0246478.
PR	14-AUG-2000; 2000US-0225213.	PR	08-NOV-2000; 2000US-0246523.
PR	14-AUG-2000; 2000US-0225214.	PR	08-NOV-2000; 2000US-0246524.
PR	14-AUG-2000; 2000US-0225266.	PR	08-NOV-2000; 2000US-0246525.
PR	14-AUG-2000; 2000US-0225267.	PR	08-NOV-2000; 2000US-0246526.
PR	14-AUG-2000; 2000US-0225268.	PR	08-NOV-2000; 2000US-0246527.
PR	14-AUG-2000; 2000US-0225270.	PR	08-NOV-2000; 2000US-0246528.
PR	14-AUG-2000; 2000US-0225447.	PR	08-NOV-2000; 2000US-0246532.
PR	14-AUG-2000; 2000US-0225757.	PR	08-NOV-2000; 2000US-0246609.
PR	14-AUG-2000; 2000US-0225758.	PR	08-NOV-2000; 2000US-0246609.
PR	14-AUG-2000; 2000US-0225759.	PR	08-NOV-2000; 2000US-0246610.
PR	18-AUG-2000; 2000US-0226279.	PR	08-NOV-2000; 2000US-0246611.
PR	22-AUG-2000; 2000US-0226681.	PR	08-NOV-2000; 2000US-0246613.
PR	22-AUG-2000; 2000US-0226688.	PR	17-NOV-2000; 2000US-0249207.
PR	22-AUG-2000; 2000US-0227182.	PR	17-NOV-2000; 2000US-0249208.
PR	23-AUG-2000; 2000US-0227009.	PR	17-NOV-2000; 2000US-0249209.
PR	30-AUG-2000; 2000US-0228924.	PR	17-NOV-2000; 2000US-0249210.
PR	01-SEP-2000; 2000US-0229287.	PR	17-NOV-2000; 2000US-0249211.
PR	01-SEP-2000; 2000US-0229343.	PR	17-NOV-2000; 2000US-0249212.
PR	01-SEP-2000; 2000US-0229344.	PR	17-NOV-2000; 2000US-0249213.
PR	01-SEP-2000; 2000US-0229345.	PR	17-NOV-2000; 2000US-0249214.
PR	05-SEP-2000; 2000US-0229509.	PR	17-NOV-2000; 2000US-0249215.
PR	05-SEP-2000; 2000US-0229513.	PR	17-NOV-2000; 2000US-0249216.
PR	06-SEP-2000; 2000US-0230437.	PR	17-NOV-2000; 2000US-0249217.
PR	06-SEP-2000; 2000US-0230438.	PR	17-NOV-2000; 2000US-0249218.
PR	08-SEP-2000; 2000US-0231242.	PR	17-NOV-2000; 2000US-0249244.
PR	08-SEP-2000; 2000US-0231243.	PR	17-NOV-2000; 2000US-0249245.
PR	08-SEP-2000; 2000US-0231244.	PR	17-NOV-2000; 2000US-0249264.
PR	08-SEP-2000; 2000US-0231413.	PR	17-NOV-2000; 2000US-0249265.
PR	08-SEP-2000; 2000US-0231414.	PR	17-NOV-2000; 2000US-0249297.
PR	08-SEP-2000; 2000US-0232080.	PR	17-NOV-2000; 2000US-0249299.
PR	12-SEP-2000; 2000US-0232081.	PR	17-NOV-2000; 2000US-0249300.
PR	12-SEP-2000; 2000US-0231968.	PR	01-DEC-2000; 2000US-0250160.
PR	14-SEP-2000; 2000US-0232397.	PR	01-DEC-2000; 2000US-0250391.
PR	14-SEP-2000; 2000US-0232398.	PR	05-DEC-2000; 2000US-0251030.
PR	14-SEP-2000; 2000US-0232399.	PR	05-DEC-2000; 2000US-0251988.
PR	14-SEP-2000; 2000US-0232400.	PR	05-DEC-2000; 2000US-0256719.
PR	14-SEP-2000; 2000US-0232401.	PR	06-DEC-2000; 2000US-0251479.
PR	14-SEP-2000; 2000US-0233063.	PR	08-DEC-2000; 2000US-0251856.
PR	14-SEP-2000; 2000US-0233064.	PR	08-DEC-2000; 2000US-0251868.
PR	14-SEP-2000; 2000US-0233065.	PR	08-DEC-2000; 2000US-0251869.
PR	21-SEP-2000; 2000US-0234223.	PR	08-DEC-2000; 2000US-0251989.
PR	21-SEP-2000; 2000US-0234274.	PR	11-DEC-2000; 2000US-0251990.
PR	25-SEP-2000; 2000US-0234997.	PR	05-JAN-2001; 2001US-0259678.
PR	25-SEP-2000; 2000US-0234998.	XX	
PR	26-SEP-2000; 2000US-0235484.	PA	(HUMA-) HUMAN GENOME SCI INC.
PR	27-SEP-2000; 2000US-0235834.	XX	
PR	27-SEP-2000; 2000US-0235836.	PI	Rosen CA, Barash SC, Ruben SM;
PR	29-SEP-2000; 2000US-0236327.	XX	
PR	29-SEP-2000; 2000US-0236367.	DR	WPI; 2001-483426/52.
PR	29-SEP-2000; 2000US-0236368.	DR	P-PSDB; AAM85891.
PR	29-SEP-2000; 2000US-0236369.	XX	

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
PS Claim 1: SEQ ID NO 3732; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX SQ Sequence 339 BP; 86 A; 84 C; 101 G; 64 T; 4 other;

Query Match 61.18; Score 174; DB 22; Length 339;
Best Local Similarity 94.98; Pred. No. 3.3e-49;
Matches 188; Conservative 1; Mismatches 8; Indels 1; Gaps 1;
Qy 1 atgcagagctctacagcctgacccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
Db 143 atgcagwtctacagctgacccagctgcccgcctatgcgaatgtgttgaggccctctgg 202
Qy 61 aggattataagaacgaggggcctatggaggccctatggaggccctgaaacgtcacagcaaca 120
Db 203 aggattataagaacgaggggcctatggaggccctatggaggccctgaaacgtcacagcaaca 261
Qy 121 ggcgcaggccctgccacgcccctttatttgcctgctacgaaagttaaaaagacattg 180
Db 262 ggcgcaggccctgccacgcccctttatttgcctgctacgaaagttaaaaagacattg 321
Qy 181 agtgatgtaataccacct 198
Db 322 agtgatgtaataccacct 339

RESULT 11
AAK52172
ID AAK52172 standard; cDNA; 1902 BP.

XX AC AAK52172;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 717.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM79039.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PT
XX
XX Claim 1: Page 2450-2452; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX SQ Sequence 1902 BP; 434 A; 532 C; 547 G; 389 T; 0 other;

Query Match 45.58; Score 129.6; DB 22; Length 1902;
Best Local Similarity 66.48; Pred. No. 9.6e-34;
Matches 186; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagctgacccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
Db 1278 atgcagagcttgcctccagatcccaagccacagtcacacaaagtgtctacgagccctcaag 1337
Qy 61 aggattataagaacgaggggcctatggaggccctatggaggccctgaaacgtcacagcaaca 120
Db 1338 aaatcatgcggacggaaggcttctggaggcccttgcgagcgctcaacgtcatgatcatg 1397
Qy 121 ggcgcaggccctgccacgcccctttatttgcctgctacgaaagttaaaaagacattg 180
Db 1398 ggtgcaggggccagcccatgcccattgtatttgcctgctatgaaacatgaaaggacttta 1457
Qy 181 agtgatgtaataccacctgggggccaatagccatattgccaatgtgtgcggccgggtgtgtg 240
Db 1458 aatgacgttttccaccaccaagaaacagccacctagccaaaggagtagctgggagtatg 1517
Qy 241 gcaacattactcatgatgcaggccatgaacctctgcgggaag 280
Db 1518 gccacctgctccacgtagcgtgaatgaatccacgagaag 1557

RESULT 12
AAF59920
ID AAF59920 standard; cDNA; 1305 BP.
XX
XX AAF59920;
XX
XX 04-MAY-2001 (first entry)
XX
XX Human mitochondrial solute carrier protein hMSC-o cDNA.

Db 316 aaatcatgcggaccgaagctctctgaggcccttgagcgctcaacgtcatgatcatg 375
Qy 121 ggcagagcctgccacgccttatttgcctgtacgaaagttaaaaaagacattg 180
Db 376 ggtcagggccagcccatgctatttctgctgtatgaaacatgaaagacattta 435
Qy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtg 240
Db 436 aatgacgtttccaccacgaagaaacagccacctagccaacggatagctgggagtatg 495
Qy 241 gcaacattactcatgatgcagcccatgaacccctgcggaag 280
Db 496 gccacctgtcccagatgcggtgaatgaatccagcagaag 535

RESULT 14
AAH90057
ID AAH90057 standard; cDNA; 2562 BP.
XX AC
XX AC
XX AAH90057;
DT 01-OCT-2001 (first entry)
XX Human bone marrow cDNA, SEQ ID NO: 301.
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
XX antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
XX immunosuppressive; gene therapy; cytokine cell proliferation;
XX cell differentiation modulator; immune disorder; infection; cancer;
XX human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX Homo sapiens.
XX OS
XX WO200153453-A2.
XX PN
XX 26-JUL-2001.
XX PD
XX XX
XX PF
XX 23-DEC-2000; 2000WO-US34960.
XX XX
XX 21-JAN-2000; 2000US-0488725.
XX PR
XX 25-APR-2000; 2000US-0552317.
XX PR
XX 09-JUL-2000; 2000US-0598042.
XX PR
XX 19-JUL-2000; 2000US-0620312.
XX PR
XX 03-AUG-2000; 2000US-0653450.
XX PR
XX 14-SEP-2000; 2000US-0662191.
XX PR
XX 19-OCT-2000; 2000US-0693036.
XX PR
XX 30-NOV-2000; 2000US-0250583.
XX XX
XX (HYSE-) HYSEQ INC.
XX PA
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Drmanac RT;
XX XX
XX WPI: 2001-488707/53.
XX DR P-PSDB; AAM00938.
XX XX
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX PT for treating e.g. cancer and immune deficiency disorders -
XX PS
XX Claim 1; Page 412-413; 648pp; English.
XX XX
XX The present sequence is one of 251 novel human polynucleotides
XX CC expressed in the bone marrow. The polynucleotide and the
XX CC polypeptide encoded by it are useful in the treatment of various
XX CC immune deficiencies and disorders. The deficiencies and disorders may
XX CC be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
XX CC infection, or may result from an autoimmune disorder, a coagulation
XX CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX CC suppression of an inflammatory response or treatment of a nervous
XX CC system disorder such as Alzheimer's disease. Detection of the presence
XX CC or increased expression of the polynucleotide or the protein it
XX CC encodes is useful for the diagnosis and/or prognosis of one

CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX XX
SQ Sequence 2562 BP; 623 A; 660 C; 627 G; 652 T; 0 other;

Query Match 44.9%; Score 128; DB 22; Length 2562;
Best Local Similarity 66.1%; Pred. No. 3.8e-33;
Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1 atgcagagttctacagctgaccagctgcgcgtctatcgcaatgtgtggagggccctctg 60
Db 2021 atgcagagtttgatccagatcccaagccagtcacacagtatctacggagccctcaag 2080
Qy 61 aggtattaaagacggagggccttatgagcccatgagggggtgaggggtcacagcaaca 120
Db 2081 aaatcatgcggaccgaagctctgagggccttgagggcctgcaacgtcatgatcatg 2140
Qy 121 ggcagggcctgccacgcctttatttgcctgtacgaaagttaaaaaagacattg 180
Db 2141 ggtcagggccgcccctgcccattgtatttgcctgtatgaaacatgaaagacattta 2200
Qy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtg 240
Db 2201 aatgacgtttccaccacgaagaaacagccacctagccaacggatagctgggagtatg 2260
Qy 241 gcaacattactcatgatgcagcccatgaacccctgcggaag 280
Db 2261 gccacctgtcccagatgcggtgaatgaatccagcagaag 2300

RESULT 15
AAZ13026
ID AAZ13026 standard; cDNA; 300 BP.
XX AC
XX AAZ13026;
XX XX
DT 12-OCT-1999 (first entry)
XX XX
XX Human gene expression product cDNA sequence SEQ ID NO:495.
XX DE
XX Human; gene; gene expression product; diagnosis; therapy; probe;
XX KW detection; mapping; tissue typing; profiling; forensic; cancer;
XX KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX XX
XX Homo sapiens.
XX OS
XX WO9938972-A2.
XX PN
XX 05-AUG-1999.
XX PD
XX 28-JAN-1999; 99WO-US01619.
XX XX
XX 03-APR-1998; 98US-0080666.
XX PR 28-JAN-1998; 98US-0072910.
XX PR 24-FEB-1998; 98US-0075954.
XX PR 31-MAR-1998; 98US-0080114.
XX PR 03-APR-1998; 98US-0080515.
XX XX
XX (CHIR) CHIRON CORP.
XX PA (HYSE-) HYSEQ INC.
XX XX
XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
XX PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
XX PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
XX PI Lanson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
XX PI Stache-Grain B, Sudduth-Klinger J, Williams LT;
XX XX
XX WPI: 1999-494092/41.
XX DR
XX Novel human genes and their expression products which are
XX PT differentially expressed in different cell types

Search completed: September 28, 2002, 08:06:28
Job time: 11583 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:55:54 ; Search time 182.06 Seconds
(without alignments)
384.519 Million cell updates/sec

Title: US-09-870-113-9

Perfect score: 285
Sequence: 1 atgcagagtctacagctga.....tgaaacctgcggaaggtga 285

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32.4	11.4	1311	4	US-09-068-140A-9
2	32.4	11.4	5150	4	US-09-068-140A-14
3	32	11.2	1643	2	US-08-933-750C-68
4	32	11.2	1643	3	US-09-234-613-68
5	29.6	10.4	4091	4	US-08-630-915A-33
6	29	10.2	1747	1	US-08-202-389-1
7	29	10.2	8600	4	US-09-457-708-1
8	28.4	10.0	4368	1	US-08-026-138E-17
9	28.4	10.0	4446	1	US-08-026-138E-6
10	28.2	9.9	7011	4	US-09-268-163-9
11	28	9.8	2379	3	US-08-537-361E-5
12	27.6	9.7	3942	4	US-09-162-484-19
13	27.2	9.5	3771	1	US-08-185-432-3
14	27.2	9.5	3771	1	US-08-185-432-23
15	27.2	9.5	3810	2	US-08-475-844-8
16	27.2	9.5	3810	5	PCT-US95-08429-8
17	27.2	9.5	5063	1	US-08-185-432-1
18	27	9.5	1710	3	US-09-106-182-1
19	27	9.5	5087	1	US-08-792-039B-3
20	27	9.5	5087	3	US-08-988-819-3
21	27	9.5	5087	3	US-09-016-534-3
22	27	9.5	7653	4	US-08-471-112A-1
23	27	9.5	7824	5	PCT-US95-06722-11
24	26.8	9.4	1192	4	US-09-142-565-1
25	26.8	9.4	1255	1	US-08-518-878B-38
26	26.8	9.4	1255	1	US-08-294-522B-38
27	26.8	9.4	1255	2	US-08-470-868A-38

28	26.8	9.4	1596	2	US-08-807-861A-38	Sequence 38, Appl
29	26.8	9.4	1596	3	US-09-210-681-38	Sequence 38, Appl
30	26.8	9.4	1596	3	US-08-946-719A-38	Sequence 38, Appl
c 31	26.8	9.4	3672	4	US-09-056-105-17	Sequence 17, Appl
32	26.8	9.4	246240	2	US-08-724-394A-20	Sequence 20, Appl
33	26.8	9.4	246240	2	US-08-724-394A-21	Sequence 21, Appl
34	26.8	9.4	246240	2	US-08-724-394A-22	Sequence 22, Appl
c 35	26.6	9.3	2130	1	US-07-952-800-1	Sequence 1, Appl
36	26.6	9.3	4692	2	US-08-916-917-1	Sequence 1, Appl
37	26.6	9.3	4692	2	US-08-972-631-1	Sequence 1, Appl
38	26.6	9.3	4692	2	US-08-972-629-1	Sequence 1, Appl
39	26.6	9.3	4692	2	US-08-972-630-1	Sequence 1, Appl
40	26.6	9.3	4692	2	US-08-672-211-1	Sequence 1, Appl
41	26.6	9.3	4692	3	US-09-225-170-1	Sequence 1, Appl
42	26.4	9.3	1777	2	US-08-937-466-5	Sequence 5, Appl
43	26.4	9.3	1777	2	US-09-172-528-5	Sequence 5, Appl
44	26.4	9.3	1777	3	US-09-318-199-5	Sequence 5, Appl
45	26.4	9.3	1777	4	US-09-503-579-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-068-140A-9
; Sequence 9, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068.140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-9

Db 482 CTACCGCTTCACTGGCACCATGGATGCCCTTCCTGTAAGATCGTGGAGCGACGAGGGCACCA 541
QY 86 ggagggcccatgaggggctgaacgtcacagcaagcgagggcctgcccacgccccttt 145
Db 542 GGACCCCTCTGGAGCGGCTCCCGCCGACCCCTGGTGATGACTGTGCCAGCTACCGCCATCT 601
QY 146 attttgcctgctacgaaaaagttaaaaaagacatctgagtgtatgcc 193
Db 602 ACTTCACTGCCTATGACCACTGAAGGCTTCCTGTGTGTCGAGGCC 649

RESULT 4
US-09-234-613-68
; Sequence 68, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYN0A01
; CLONE: 724157

US-09-234-613-68
Query Match 11.2%; Score 32; DB 3; Length 1643;
Best Local Similarity 49.4%; Pred. No. 0.16;
Matches 83; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 26 ctgcccgcctatcgcaatgtgttgaggccctctgaggattataagaacgagggcctat 85
Db 482 CTACCGCTTCACTGGCACCATGGATGCCCTTCCTGTAAGATCGTGGAGCGACGAGGGCACCA 541

QY 86 ggagggcccatgaggggctgaacgtcacagcaagcgagggcctgcccacgccccttt 145
Db 542 GGACCCCTCTGGAGCGGCTCCCGCCGACCCCTGGTGATGACTGTGCCAGCTACCGCCATCT 601
QY 146 attttgcctgctacgaaaaagttaaaaaagacatctgagtgtatgcc 193
Db 602 ACTTCACTGCCTATGACCACTGAAGGCTTCCTGTGTGTCGAGGCC 649

RESULT 5
US-08-630-915A-33
; Sequence 33, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4091 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-630-915A-33

Query Match 10.4%; Score 29.6; DB 4; Length 4091;
Best Local Similarity 51.5%; Pred. No. 1.7;
Matches 68; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 10 ctacagcctgaccagctgctccgctatcgcaatgtgttgaggccctctgaggattata 69
Db 995 CGAGAACTGGAGCAGGCCCATCCGGGGCCGATGCCAGGAGACCTCAGGTGGTTCCGC 1054
QY 70 agaacgaggggcctatggagggcccatgaggggctgaacgtcacagcaacagggcgaggg 129
Db 1055 AGCACAGTGGCCCGGGATGCCATGATGCCCACTGACTGACTGACTGACTGACTGACTG 1114
QY 130 cctgcccacgccc 141
Db 1115 CTCCCGCACACC 1126

```

RESULT      6
US-08-202-389-1
; Sequence 1, Application US/08202389
; Patent No. 5536636
; GENERAL INFORMATION:
; APPLICANT: Freeman Jr., Robert M.
; APPLICANT: Plutzky, Jorge
; APPLICANT: Neel, Benjamin G.
; APPLICANT: Rosenberg, Robert D.
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
; TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,389
; FILING DATE: 28-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,926
; FILING DATE: 01-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,141
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/721,112
; FILING DATE: 26-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: BIH92-05MA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1540
; US-08-202-389-1

```

	Query Match	10.2%	Score 29;	DB 1;	Length 1747;
	Best Local Similarity	57.0%;	Pred. No. 1.8;		
	Matches 53;	Conservative 0;	Mismatches 40;	Indels 0;	Gaps 0;
Qy	96	gaggggctgaacgtcacagcaacagcgaggcgctgcacacgcccccttttattttgctg	155		
Db	105	GAGGCGAGGCAGAGTACTGCTGCAGGCCAAGGCGAGCCCTGGACATTTCTTGTGCGTG	164		
Qy	156	ctacgaaaagttaaanaagacattgagtgtg	188		
Db	165	AGAGTCTTCAGCCCAACCTGGTGATTTTGTGCTCT	197		

RESULT 7
US-09-457-708-1/c

```

; Sequence 1, Application US/09457708
; Patent No. 6326483
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, David J.
; APPLICANT: Sampson, Julian R.
; APPLICANT: Povey, Sue
; APPLICANT: van Slegtenhorst, Marjon
; APPLICANT: Halley, Dicky
; TITLE OF INVENTION: Compositions and Methods Based U
; TITLE OF INVENTION: Sclerosis-1 (TSC-1) Gene and Gene Product
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Vinson & Elkins
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20004-1008
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/457,708
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanzo, Michael A.
; REGISTRATION NUMBER: 36,912
; REFERENCE/DOCKET NUMBER: BRI331/42002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 639-6585
; TELEFAX: (202) 639-6604
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-09-457-708-1

Query Match 10.2%; Score 29; DB 4; Length 8600;
Best Local Similarity 58.8%; Pred.No.3.8;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps

QY 100 ggctgaacgtcacagcaacgagcgacggcctgccacgccccttatttgctgtac 159
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7750 GGACAAAGGTCACAGCACACGCCATAGGGGGCCAGTCCTCGACCTAAACTTGTCTTTCAC 7691
QY 160 gaaaagttaaaaaagacattgagt 184
|| || |||| || |||||
Db 7690 CTACACACAAACCTTAATCAAGTG 7666

RESULT 8
US-08-026-138E-17/c
; Sequence 17, Application US/08026138E
; Patent No. 5502166
; GENERAL INFORMATION:
; APPLICANT: Masayoshi MISHINA
; TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nishiohata Residence 1-107
; STREET: 5214, Nishiohata-machi
; CITY: Niigata-shi
; STATE: Niigata-ken
; COUNTRY: JAPAN

```



```

Db 6357 TGAAGAGAGGAGGAGGCTCCCTCCGCTCTGGACCGGCTCGCTCTTACGCTC 6298
      ||| |||| ||| ||| ||| |||| | |||| | |||| | |||
Qy 227 cggcgggtgtgtggaacattcatctgatgcagccatgaacc 271
      ||||| | ||| | ||| |||| |||| | |||
Db 6297 CGCGGGCAGGCTGTGGACCTTCTCCATGGGGCAGGCCAGATCC 6253
      ||||| | ||| | ||| |||| |||| | |||

RESULT 11
US-08-537-361E-5/C
; Sequence 5, Application US/08537361E
; Patent No. 6121037
; GENERAL INFORMATION:
; APPLICANT: Stojiljkovic, Igor
; APPLICANT: So, Magdalene
; APPLICANT: Hwa, Vivian
; APPLICANT: Heffron, Fred
; APPLICANT: Nassif, Xavier
; TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor
; TITLE OF INVENTION: Genes and Uses
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,361E
; FILING DATE: 02-OCT-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6121037nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 94,784-A
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2379 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2376
US-08-537-361E-5

Query Match 9.8%; Score 28; DB 3; Length 2379;
Best Local Similarity 46.4%; Pred. No. 4.6; Mismatches 105; Indels 0; Gaps 0;
Matches 91; Conservative 0;

Qy 87 gagcccatgaggggctgaacgtcacagcaacagcgagggcctgccacgcgccttta 146
      |||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1390 GAGGCGTCATTGTGTATGATCGTAAACGATACCTGCGCGCTACTGAACACGTCGTCC 1331
      |||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 147 ttttgcctgtcacgaaagtttaaaagacattgagtgatgaatccaccctgggggcaa 206
      || | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1330 ATTGAATTGGTTCAGACATGAGAACCGTAGTGTGTTTTCACCGGATGCTGGATAC 1271
      || | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 207 tagccatattgccaatggtgcggccgggtgtgtggcaacattacttcatgatgcagcoat 266
      || | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1270 TGCTGGTGGTTGGAACAAACACGCGCGCTGAAGTAGTAATAGTCGCGGTTTAGGTTTCAA 1211
      || | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Qy 267 gaaccctgcggaaggc 282
      | | | | | | | | | | | | | | | | | |
Db 1210 AATCAGCGCGCTGCC 1195
      | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-162-484-19
; Sequence 19, Application US/09162484
; Patent No. 6248724
; GENERAL INFORMATION:
; APPLICANT: Phillips, M. Ian
; APPLICANT: Mohuczy, Dagmara
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
; FILE REFERENCE: UFLA:087/UFLA087P
; CURRENT APPLICATION NUMBER: US/09/162,484
; CURRENT FILING DATE: 1998-09-25
; EARLIER APPLICATION NUMBER: 60/059,661
; EARLIER FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 3942
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-162-484-19

Query Match 9.7%; Score 27.6; DB 4; Length 3942;
Best Local Similarity 52.6%; Pred. No. 8.2; Mismatches 54; Indels 0; Gaps 0;
Matches 60; Conservative 0;

Qy 133 gccacagccctttatttgcctgtctacgaaaagttaaaaaagacattgagtgtgtaatc 192
      | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db 2989 gccacatccagttatttcacgtacacaaagacttgccttgaccttcggaggggcgcc 3048
      | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Qy 193 caccctgggggcaatagccattatggccaatggtgcggccgggtgtgtggcaaca 246
      | |||| | || | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db 3049 aaccccggttttcacgtatgaggtatgttggctctgtctgtctatca 3102
      | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||

RESULT 13
US-08-185-432-3/C
; Sequence 3, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872

```

REFERENCE/DOCKET NUMBER: 7326-006
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3771 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 345..2558
US-08-185-432-3

Query Match 9.5%; Score 27.2; DB 1; Length 3771;
Best Local Similarity 49.3%; Pred. No. 11;
Matches 71; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 9 tctacagctgaccagctgccgctatcgcaatgtgttgaggccctctgaggattat 68
Db 1035 TGTACTGGCTGCTCAGCTGGCGGAATTCGACTTGAGTTGGTGGCCTGTGTCGCGTTA 976
Qy 69 aagaacggaggccctatgaggcccatgaggggcgtgaacgtcacagcaacagggcgcagg 128
Db 975 GTTTCACCAAGGATACGGCGCTGTGGTACGCCGAATGCTGGCGCATGGTCCGCTGG 916
Qy 129 gcctgcccacgcccctttatttgc 152
Db 915 GTTGGCGCACGTGGGTGAGATTGC 892

RESULT 14
US-08-185-432-23
Sequence 23, Application US/08185432
Patent No. 5750652
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Busseau, Isabelle
APPLICANT: Diederich, Robert J.
APPLICANT: Xu, Tian
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:
LENGTH: 3771 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-185-432-23

Query Match 9.5%; Score 27.2; DB 1; Length 3771;
Best Local Similarity 49.3%; Pred. No. 11;
Matches 71; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 9 tctacagctgaccagctgccgctatcgcaatgtgttgaggccctctgaggattat 68
Db 2737 TGTACTGGCTGCTCAGCTGGCGGAATTCGACTTGAGTTGGTGGCCTGTGTCGCGTTA 2796
Qy 69 aagaacggaggccctatgaggcccatgaggggcgtgaacgtcacagcaacagggcgcagg 128
Db 2797 GTTTCACCAAGGATACGGCGCTGTGGTACGCCGAATGCTGGCGCATGGTCCGCTGG 2856
Qy 129 gcctgcccacgcccctttatttgc 152
Db 2857 GTTGGCGCACGTGGGTGAGATTGC 2880

RESULT 15
US-08-475-844-8
Sequence 8, Application US/08475844
Patent No. 5972643
GENERAL INFORMATION:
APPLICANT: Lobanenko, Victor V.
APPLICANT: Neiman, Paul E.
APPLICANT: Klenova, Elena M.
APPLICANT: Goodwin, Graham H.
APPLICANT: Filippova, Galina N.
APPLICANT: Collins, Steven J.
TITLE OF INVENTION: CTCF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,844
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,680
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14538A-11-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3810 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA

```
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: human CTCF cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 292..2475
; FEATURE:
; NAME/KEY: exon
; LOCATION: 281..1074
; OTHER INFORMATION: /label= exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1075..1245
; OTHER INFORMATION: /label= exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1246..1379
; OTHER INFORMATION: /label= exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1380..1499
; OTHER INFORMATION: /label= exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1500..1649
; OTHER INFORMATION: /label= exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1650..1810
; OTHER INFORMATION: /label= exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1810..1992
; OTHER INFORMATION: /label= exon8
;
US-08-475-844-8
```

```
Query Match          9.5%; Score 27.2; DB 2; Length 3810;
Best Local Similarity 64.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 45 gttgaggccctctgaggattataagaacggaggccctatggagcccatgaggggct 104
    |||||
Db 682 GTACAAGAAC TTCAGGGGCTTATGAAATGAAGTCTCTAAAGAGGGCCTTGCGGAAAGT 741
    |||||

Qy 105 gaac 108
    |||||
Db 742 GAAC 745
```

Search completed: September 28, 2002, 07:56:00
Job time: 11245 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 06:22:48 ; Search time 4881.77 Seconds
(without alignments)
787.959 Million cell updates/sec

Title: US-09-870-113-9

Perfect score: 285

Sequence: 1 atcagaggtctacagcctga.....tgaacctgcggaaggtga 285

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	280	98.2	747	10	BI827322 603077989
2	277.6	97.4	888	9	AL530804 AL530804
3	265.8	93.3	480	10	BI045863
4	257.4	90.3	573	10	BI541703 455541 MA
5	246.8	86.6	598	10	BE913718 601669072
6	246.8	86.6	645	10	BG964218 602828996
7	246.8	86.6	1131	10	BG295496 602392826
8	245.2	85.0	560	10	BI794616 ic92e02.y
9	244.6	85.8	574	10	BI185580 UNL-P-FN-
10	234	82.1	501	9	AW211366 uc80b07.y
11	224.4	78.7	701	10	BI831943 603076873
12	223.4	78.4	336	9	AW326482 19106 MAR
13	223.2	78.3	817	10	BI103329 602888838
14	195.4	68.6	546	10	BM488747 pgn2n.pko
15	191.8	67.3	855	10	BI412175 602965377
16	191.6	67.2	453	10	BF916224 CM2-UT011
17	188.4	66.1	967	10	BI739710 603361983

18	185.6	65.1	397	9	BE012485
19	182.2	63.9	697	10	BI854643
20	168	58.9	722	10	BI876051
21	166.2	58.3	622	10	BJ039504
22	166	58.2	681	9	AV704087
23	156.8	55.0	421	10	BI448355
24	149.8	52.6	637	9	AL654963
25	143.8	50.5	417	9	AA104365
26	142.6	50.0	546	9	BE014800
27	140.8	49.4	561	9	AI942584
28	139.4	48.9	678	9	BB654057
29	139.2	48.8	1295	11	AK015770
30	137.8	48.4	916	10	BI757180
31	129.6	45.5	567	10	BJ035232
32	129.6	45.5	612	10	BJ012200
33	128.2	45.0	512	9	AW210205
34	128	44.9	684	9	AI133696
35	128	44.9	750	10	BI831750
36	126.4	44.4	488	9	AI310713
37	126	44.2	309	9	AI630395
38	125.4	44.0	729	10	BI460337
39	116	40.7	803	10	BG777404
40	115.2	40.4	1209	11	AK006155
41	113.8	39.9	331	10	BI449976
42	109.8	38.5	666	9	BB644566
43	107.6	37.8	593	10	BG895490
44	105.2	36.9	483	9	AA061624
45	103.4	36.3	350	12	AZ693917

ALIGNMENTS

RESULT 1

BI827322

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI827322 747 bp mRNA linear EST 04-OCT-2001
603077989F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169694 5',
mRNA sequence.
BI827322 GI:15938872
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 747)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11422 row: a column: 23
High quality sequence stop: 743.
Location/Qualifiers
1. .747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169694"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range

FEATURES
source

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT 188 a 206 c 201 g 151 t 1 others
ORIGIN

Query Match 98.2%; Score 280; DB 10; Length 747;
Best Local Similarity 100.0%; Pred. No. 3.1e-76;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctctacagcctgagccagctgcccgcctatcgcaatgtgttgagagccctctgg 60
|||||
Db 84 ATGCAGAGTCTACAGCCTGACCCAGCTGCCGCTATGCAATGTGTGGAGGCCCTCTGG 143
|||||
QY 61 aggtattataagaacgagggccctatggagggccatgaggggctgaaagctcacagcaaca 120
|||||
Db 144 AGGATTATAAGAAACGAGGGCCCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 203
|||||
QY 121 ggcgagggcctgccacagcccttatttgcctgctctacgaaagttaaaaaagacattg 180
|||||
Db 204 GCGCAGGGCCCTGCCACGCCCCCTTTATTTGCTGCTACGAAAAGTTAAAAAGACATTG 263
|||||
QY 181 agtgatgtaataccaccctgggggcaatagccatattgccaatggtgcggccgggtgtgtg 240
|||||
Db 264 AGTGATGTAATCCACCCTGGGGGCAATAGCCATATGCCATGTGGCGCCGGGTGTGTG 323
|||||

QY 241 gcaacattactctatgctgagccatgaacctgcggaag 280
|||||
Db 324 GCAACATTACTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 363
|||||

RESULT 2

AL530804
LOCUS AL530804 LTI_NFL001_NBC4 888 bp mRNA linear EST 13-FEB-2001
DEFINITION prime, mRNA sequence.
ACCESSION AL530804
VERSION AL530804.1 GI:12794297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 888)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1 . 888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD008TH17"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; vector: pCWSPT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCWSPT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

FEATURES

source

ORIGIN

Query Match 97.4%; Score 277.6; DB 9; Length 888;
Best Local Similarity 99.3%; Pred. No. 1.8e-75;
Matches 278; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgcagagctctacagcctgagccagctgcccgcctatcgcaatgtgttgagagccctctgg 60
|||||
Db 417 ATGCAGAGTCTACAGCCTGACCCAGCTGCCGCTATGCAATGTGTGGAGGCCCTCTGG 476
|||||
QY 61 aggtattataagaacgagggccctatggagggccatgaggggctgaaagctcacagcaaca 120
|||||
Db 477 AGGATTATAAGAAACGAGGGCCCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 536
|||||
QY 121 ggcgagggcctgccacagcccttatttgcctgctctacgaaagttaaaaaagacattg 180
|||||
Db 537 GCGCAGGGCCCTGCCACGCCCCCTTTATTTGCTGCTGCBACGAAAAGTTAAAAAGACATTG 596
|||||
QY 181 agtgatgtaataccaccctgggggcaatagccatattgccaatggtgcggccgggtgtgtg 240
|||||
Db 597 AGTGATGTTATCCACCCTGGGGGCAATAGCCATATGGCAATGTGGCGCCGGGTGTGTG 656
|||||
QY 241 gcaacattactctatgctgagccatgaacctgcggaag 280
|||||
Db 657 GCAACATTACTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 696
|||||

RESULT 3

BI045863
LOCUS BI045863 MR3-FN0209-300101-004-h10 FN0209 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BI045863
VERSION BI045863.1 GI:14452485
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 480)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hate M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&t2=MR3-FN0209-300101-004-h10&t3=2001-01-30&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 480.
Location/Qualifiers
1 . 480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0209"
/dev_stage="Adult"

TITLE

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&t2=MR3-FN0209-300101-004-h10&t3=2001-01-30&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 480.

Location/Qualifiers

1 . 480

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="FN0209"

/dev_stage="Adult"

; Site.2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent

FEATURES

source

1 . 480

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="FN0209"

/dev_stage="Adult"

; Site.2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent

products derived from ORESTES PCR (U.S. Letters Patent

products derived from ORESTES PCR (U.S. Letters Patent

products derived from ORESTES PCR (U.S. Letters Patent

products derived from ORESTES PCR (U.S. Letters Patent

products derived from ORESTES PCR (U.S. Letters Patent

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 116 a 126 c 146 g 92 t

ORIGIN

Query Match 93.3%; Score 265.8; DB 10; Length 480;
Best Local Similarity 98.9%; Pred. No. 7.2e-72;
Matches 278; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 atgcagagctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
Db 74 ATGCAGAGCTACAGCTGACCCAGCTGCCCGCTATCGCAATGTGTGGAGGCCCTCTGG 133
|||||

QY 61 aggtattataagacgagggcctatgagcccatgagggcctgaacgtcacagcaaca 120
|||||
Db 134 AGGATTATTAAGAACGAGGGCCCTATGGAGCCCATGAGGGCGTGAACGTGCACAGCAACA 193
|||||

QY 121 ggcgcaggcctgcccacgcctcttatttgcctgctacgaaagttaaaaaagacattg 180
|||||
Db 194 GCGCAGGGCCTGCCACACGCGCTTTATTTTGCCTGTACGAAAGTTAAAAAGACATTG 253
|||||

QY 181 agtgatgtaataccacct-gggggcaatagcattgtgcaatgtgtgcggcggtgtgt 239
|||||

Db 254 AGTGATGTAATCCACCTCGGGGGCAATAGCCATATGTGCAATGTGTGCGCGCGGTGTGT 313
|||||

QY 240 ggcacattactctcatgatgcagccatgaacctgcggaag 280
|||||

Db 314 GGCACCATTTACTTCATGATGACCCCATGAACCTGCGGAAG 354
|||||

RESULT 4
BI541703 573 bp mRNA linear EST 30-AUG-2001
LOCUS 455441 MARC IBOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BI541703
ACCESSION BI541703.1 GI:15382815
VERSION
KEYWORDS EST.
SOURCE Bos taurus
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrnkruug,S.C., Bennett ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Contact: Smith TPL
MEDLINE USDA, ARS, US Meat Animal Research Center
COMMENT PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCAGCG
Plate: 121 row: O column: 24
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1..573
/organism="Bos taurus"
/db_xref="taxon:9913"

/clone_lib="MARC IBOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

BASE COUNT 121 a 166 c 176 g 110 t

ORIGIN

Query Match 90.3%; Score 257.4; DB 10; Length 573;
Best Local Similarity 94.3%; Pred. No. 3.1e-69;
Matches 267; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 atgcagagctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
Db 174 ATGCAGAGCTACAGCGGATCCAGCGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 233
|||||

QY 61 aggtattataagacgagggcctatgagcccatgagggcctgaacgtcacagcaaca 120
|||||
Db 234 AGGATTATTAAGAACGAGGGCCTGTGGAGGCCCATCGAGGCCGTGAACCTCAGCAACA 293
|||||

QY 121 ggcgcaggcctgcccacgcctcttatttgcctgctacgaaagttaaaaaagacattg 180
|||||
Db 294 GCGCAGGGCCTGCCACACGCGCTCTATTTTGCCTGTATGAAAAGTTAAAAAGACATTG 353
|||||

QY 181 agtgatgtaataccacctgggggcaatagcattgtgcaatgtgtgcggcggtgtgtg 240
|||||

Db 354 AGTGATGTAATCCACCTCGGGGCAATAGCCATATGTGCAATGTGTGCGCGGTGTGTG 413
|||||

QY 241 gcaacattactctcatgatgcagccatgaacctgcggaaggt 283
|||||

Db 414 GCAACATTACTTCATGATGCTGCCATGAATCCAGTGGAAAGACT 456
|||||

RESULT 5
BE913718 598 bp mRNA linear EST 29-SEP-2000
LOCUS 601669072F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5',
DEFINITION BE913718.1 GI:10411618
ACCESSION BE913718
VERSION BE913718.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9146 row: O column: 17
High quality sequence stop: 593.
Location/Qualifiers
1..598
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969256"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 141 a 175 c 172 g 110 t

ORIGIN

Query Match 86.6%; Score 246.8; DB 10; Length 598;
Best Local Similarity 92.2%; Pred. No. 6.2e-66;
Matches 260; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 104 ATGCAGAGCCTACAGCCTGACCCAGCCGCCGCTATCGAAGCTGTGGAGGCTCTCTGG 163

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 61 aggattataaagagagggcctatggagccatgaagggggtgaacgtcacagcaaca 120

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 164 AGAATCATGAGACAGAGGGCCGTGTGGAGGCCATGCGGGGCTCAAGCTCACAGCAACA 223

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 121 ggcgcagggcctgcccagcgcctttatttgcctgctacgaaaagttaaaaagacattg 180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 224 GCGCGGGCCTGCCACGCCCTCTATTGTGCTCTACGAAAGCTTAAAAAGACATTG 283

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 181 agtcatgtaatccaccctgggggcaatagccatatgtccaatggtgcggccgggtgtgtg 240

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 284 AGTGACGTAATCCACCCAGGGGCAATAGCCATATTGCCAATGTGTCAGCGCGGATGTGTG 343

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 241 gcaacattacttcagtcagccatgaacctgcggaaggc 282

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 344 GCGACATTACTTCATGATGCAGCCATGAATCCAGCGGAAGTC 385

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 6

BC964218 BC964218 645 bp mRNA linear EST 12-JUN-2001

LOCUS 602828996F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4983815 5',

DEFINITION mRNA sequence.

ACCESSION BC964218

VERSION BC964218.1 GI:14351855

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 645)

JOURNAL NIH-MGC http://mgc.nci.nih.gov/

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10988 row: p column: 24

High quality sequence start: 2

High quality sequence stop: 632.

Location/Qualifiers

1. .645

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4983815"

/clone_lib="NCI_CGAP_Co24"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.6 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 147 a 189 c 193 g 116 t

ORIGIN

Query Match 86.6%; Score 246.8; DB 10; Length 645;
Best Local Similarity 92.2%; Pred. No. 6.3e-66;
Matches 260; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 173 ATGCAGAGCCTACAGCCTGACCCAGCCGCCGCTATCGAAGCTGTGGAGGCTCTCTGG 232

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 61 aggattataaagagagggcctatggagcccatgaagggggtgaacgtcacagcaaca 120

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 233 AGAATCATGAGACAGAGGGCCTGTGGAGGCCATGCGGGGCTCAAGCTCACAGCAACA 292

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 121 ggcgcagggcctgcccagcgcctttatttgcctgctacgaaaagttaaaaagacattg 180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 293 GCGCGGGCCTGCCACGCCCTCTATTGTGCTCTACGAAAGCTTAAAAAGACATTG 352

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 181 agtcatgtaatccaccctgggggcaatagccatatgtccaatggtgcggccgggtgtgtg 240

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 353 AGTGACGTAATCCACCCAGGGGCAATAGCCATATTGCCAATGTGTCAGCGCGGATGTGTG 412

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 241 gcaacattacttcagtcagccatgaacctgcggaaggc 282

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 413 GCGACATTACTTCATGATGCAGCCATGAATCCAGCGGAAGTC 454

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 7

BC295496 BC295496 1131 bp mRNA linear EST 21-FEB-2001

LOCUS 602392826F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4504638 5',

DEFINITION mRNA sequence.

ACCESSION BC295496

VERSION BC295496.1 GI:13057189

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 1131)

JOURNAL NIH-MGC http://mgc.nci.nih.gov/

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10377 row: c column: 07

High quality sequence stop: 753.

Location/Qualifiers

1. .1131

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:4504638"

/clone_lib="NIH_MGC_94"

/tissue_type="retina"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-dt primed.

Average insert size 3.3 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 261 a 335 c 375 g 160 t

ORIGIN

Query Match 86.6%; Score 246.8; DB 10; Length 1131;

```
Best Local Similarity 92.2%; Pred. No. 7.3e-66;
Matches 260; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 atgcagagtctacagctgaccagctgcccgtctatcgcaatgtgttgagccctctgg 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 261 ATGCAGAGCTACAGCTGACCCAGCGCCGCTATCGGAAGCTGTGGAGGCTCTCTGG 320
QY 61 aggattataagaacagagggccctatgagggcccatgaggggctgaacgtcacagaaca 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 321 AGAATCATGAGACAGAGGCGCTGTGGAGGCCATCGCGGGGCTGAACGTCACAGCAACA 380
QY 121 ggcgaggcctgcccagccctttatttctgctgctacgaaagttaaagaacattg 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 381 GCGCGGGGCGCTGCCACGCCCTCTATTTTGCTGCTACGAAAGTTAAAAAGACATTG 440
QY 181 agtgtgttaatccaccctggggcctatgacccatattgccaatggtgcgcccgtgtgtg 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 441 AGTACGTAAATCCAGCGGGGCAATACCCATATTTGCCAATGTGGCAGCCGGAATGTGTG 500

QY 241 gcaacattactctatgatgcagccatgaacctgcggaagc 282
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 501 GCGACATTACTTTCATGATGCAGCCATGAATCAGCGGAAGTC 542

RESULT 8
BI794616
LOCUS
DEFINITION
ic92e02.y1 Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus
musculus cDNA 5' similar to TR:Q23125 Q23125 W02B12.9 PROTEIN. [1]
; mRNA sequence.
ACCESSION
BI794616
VERSION
BI794616.1 GI:15822341
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 560)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,I.,
Jackson,I. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: ic92e02.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
Seq primer: 40RP from Gibco
High quality sequence stop: 430.
Location/Qualifiers
1..560
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1
NI-MMS1"
/sex="Both for embryonic & newborn, male for adult and
adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
adult, mixed"
/lab_host="DH10B"

/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; Five
libraries representing E10.5/12.5 pancreatic bud, E16.5
pancreas, newborn pancreas, adult pancreas, and adult
islets of Langerhans were separately constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
was made by oligo-dT priming and size-selected by column
fractionation. Libraries were amplified once on solid
support and plasmid DNA from each library was prepared
and mixed in equal amounts. The mixed library DNA was
normalized by method #4 from Bonaldo, Lennon, and Soares
1996 Genome Research 6:791-806; 0.5 microgram
single-stranded mixed library plasmid DNA was mixed with
5 micrograms PCR product representing mixed library
(inserts and hybridized to an EcoT of 6. Single-stranded
(unhybridized) plasmids were isolated by hydroxyapatite
chromatography and used to make this library."

BASE COUNT 134 a 158 c 163 g 105 t
ORIGIN

Query Match 86.0%; Score 245.2; DB 10; Length 560;
Best Local Similarity 91.8%; Pred. No. 1.9e-65;
Matches 259; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 atgcagagtctacagctgaccagctgcccgtctatcgcaatgtgttgagccctctgg 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84 ATGCAGAGCTACAGCTGACCCAGCGCCGCTATCGGAAGCTGTGGAGGCTCTCTGG 143
QY 61 aggattataagaacagagggccctatgagggcccatgaggggctgaacgtcacagaaca 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 144 AGAATCATGAGACAGAGGCGCTGTGGAGGCCATCGCGGGGCTGAACGTCACAGTAACA 203
QY 121 ggcgaggcctgcccagccctttatttctgctgctacgaaagttaaagaacattg 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 204 GCGCGGGGCGCTGCCACGCCCTCTATTTTGCTGCTAGAAAAAGTTAAAAAGACATTG 263
QY 181 agtgtgttaatccaccctgggggcaatagccatattgccaatggtgcgcccgtgtgtg 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 264 AGTACGTAAATCCAGCGGGGCAATACCCATATTTGCCAATGTGGCAGCCGGAATGTGTG 323
QY 241 gcaacattactctatgatgcagccatgaacctgcggaagc 282
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 324 GCGACATTACTTTCATGATGCAGCCATGAATCAGCGGAAGTC 365

RESULT 9
BI185580
LOCUS
DEFINITION
UNL-P-FN-cy-f-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
UNL-P-FN-cy-f-02-0-UNL 3', mRNA sequence.
ACCESSION
BI185580
VERSION
BI185580.1 GI:14659989
KEYWORDS
EST.
SOURCE
pig.
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 574)
Caetano,A.R., Johnson,R.K. and Pomp,D.
Generation and sequence characterization of a normalized cDNA
library from swine ovarian follicles
Unpublished (2001)
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message.
Seq primer: M13 -29
POLYA=No.
```


CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L14M1419 row: a column: 20
 High quality sequence start: 29
 High quality sequence stop: 699.
 Location/Qualifiers
 1..701
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5168539"
 /clone_lib="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dr primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."
 BASE COUNT 176 a 167 c 208 g 150 t
 ORIGIN

Query Match 78.7%; Score 224.4; DB 10; Length 701;
 Best Local Similarity 99.6%; Pred. NO. 5.9e-59;
 Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 atgcagagctctacagcctgaccagctccgcctatcgcaatgtgttgaggccctctgg 60
 Db 302 ATGCAGAGCTCTACAGCTTACAGCCAGCTGCCAGCTATGCCAATGTGTGAGGCCCTCTGG 361
 QY 61 agattataaagacgagggccctatggagggccatgaggggctgaagcgtcacagcaaca 120
 Db 362 AGGATTATAAAGAACGGAGGCCCTATGGAGGCCCATGGGGCTCAACGTCACAGCAACA 421
 QY 121 ggcgagggcctgcccagcccttattttgctctcagaaaagttaaaagacattg 180
 Db 422 GCGCAGGGCCCTGCCAGCCCTTTATTTCCTCTCTAGCAAGTTAAAAAGACATTG 481
 QY 181 agtgatgtaatccacctgggggccaatagccatattgccaatgggtg 226
 Db 482 AGTGATGTAATCCACCCTGGGGGCAATAGCCATATTGCCATGCTG 527

RESULT 12
 AW326482 336 bp mRNA linear EST 25-APR-2001
 LOCUS
 DEFINITION 19106 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION AW326482
 VERSION AW326482.1 GI:6762403
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 336)
 Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
 ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.,
 Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
 Keefe,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013

COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCAGCAG
 Plate: 8 row: F column: 16
 Seq primer: ATTAGTGCACACTATAG.

FEATURES

Location/Qualifiers
 1..336
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 1BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."

BASE COUNT 72 a 96 c 107 g 61 t
 ORIGIN

Query Match 78.4%; Score 223.4; DB 9; Length 336;
 Best Local Similarity 95.4%; Pred. NO. 9.8e-59;
 Matches 230; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 1 atgcagagctctacagcctgaccagctccgcctatcgcaatgtgttgaggccctctgg 60
 Db 96 ATGCAGAGCTCTACAGCCGATCCAGCCGCCCTATGCCAATGTGTGAGGCCCTCTTGG 155
 QY 61 agattataaagacgagggccctatggagggccatgaggggctgaagcgtcacagcaaca 120
 Db 156 AGGATTATAAAGAACGGAGGCCCTGTGGAGGCCATGCGAGGCCCTCAACGTCACAGCAACA 215
 QY 121 ggcgagggcctgcccagcccttattttgctctcagaaaagttaaaagacattg 180
 Db 216 GCGCAGGGCCCTGCCAGCCCTCTATTTCCTCTCTATGAAAGTTAAAAAGACATTG 275
 QY 181 agtgatgtaatccacctgggggccaatagccatattgccaatgggtg 240
 Db 276 AGTGATGTAATCCACCCTGGGGGCAATAGCCATATTGCCATGTCGGCTGGTGTG 335
 QY 241 g 241
 Db 336 G 336

RESULT 13
 B1103329

LOCUS
 DEFINITION 602888838F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5043992
 5', mRNA sequence.

ACCESSION B1103329
 VERSION B1103329.1 GI:14554222
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 817)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: MGC Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LUAM1120 row: 1 column: 09
 High quality sequence stop: 777.
 Location/Qualifiers

1. .817
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:5043992"
 /clone_lib="NCI_CGAP_Kid14"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library. |"

FEATURES

source

1. .546
 /organism="Gallus gallus"
 /strain="Commercial broiler and Ottawa Res. Centre
 Strains 90 & 21"
 /db_xref="taxon:9031"
 /clone="pgm2n.pk008.l2"
 /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
 and Epiphyseal Growth Plate cDNA library (pgm2n)"
 /sex="Male and Female"
 /tissue_type="Breast muscle, leg muscle and epiphyseal
 growth plate"
 /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
 ,11 weeks);growth plate(1d,7d,14d post-hatch)"
 /lab_host="E. coli EMDH10B"
 /note="Vector: pCMVSPORT6; Library made from equivalent
 pools of total RNA isolated from each tissue (embryonic
 muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
 plate 33.3% of the final RNA pool). Single pass sequencing
 from 5'-end"

BASE COUNT 147 a 222 c 316 g 132 t

ORIGIN

Query Match 78.3%; Score 223.2; DB 10; Length 817;
 Best Local Similarity 91.2%; Pred. No. 1.5e-58;
 Matches 259; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

Qy 1 atcgagagtctacagcgtgacccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
 Db 422 ATGCAGAGCCTACAGCCTGACCCAGCCGCCGCTATCGGAACGTTTGGAGGCTCTGG 481
 Qy 61 aggattataagaacggaggcctatgagggccatgaggggctgaacgtcacagcaaca 120
 Db 482 AGAATCAGGACAGAGAGGGCCTGTGAGGCGCCATGCGGGGGCTGACGCTCAGACCA 541
 Qy 121 ggcagaggcctgcccagcccttatttgcctgctacgaaaaagttaaaagacattg 180
 Db 542 GCGCGGGGCGCTGCCAGCCCTCTATTTCCTGCTACGAAAAAGTTAAAAAAGACATTG 601
 Qy 181 agtgatgaattccacctgggggcaatagccattatcccaatg-tacggccgggtgtgt 239
 Db 602 AGTACGTAATCCACCCAGGGGGCAATAGCCATATGCCAATGTTGAGCCGGATGTGT 661
 Qy 240 ggcacattac-ttcattgatgcagccatgaacctgcggaaggc 282
 Db 662 GCGACATTTCTTCATGATGCAGGCATGAATCCAGCGGAAGTC 705

RESULT 14

BM488747 546 bp mRNA linear EST 07-FEB-2002
 LOCUS pgm2n.pk008.l2 Normalized Chicken Breast Muscle, Leg Muscle, and
 DEFINITION Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
 clone pgm2n.pk008.l2 5' similar to ref|XP_050766.1 (XM_050766)
 putative mitochondrial solute carrier (Homo sapiens)
 gb|AAK49519.1|AF327402.1 (AF327402) putative mitochondrial solute
 carrier splice variant [Homo sapiens], mRNA sequence.

ACCESSION BM488747
 VERSION BM488747.1 GI:18609678

KEYWORDS EST.

SOURCE chicken.

ORGANISM

Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE

1 (bases 1 to 546)
 Cogburn, L.A. and Monson-Orran, E.
 ESTs from Normalized Chicken Breast
 Epiphyseal Growth Plate cDNA library, USDA/IRAFs Animal Genome
 Project

JOURNAL

Unpublished (2002)
 Contact: Larry A. Cogburn
 University of Delaware

Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2822
 Email: cogburn@udel.edu, www.chickest.udel.edu.

Location/Qualifiers

source

1. .546
 /organism="Gallus gallus"
 /strain="Commercial broiler and Ottawa Res. Centre
 Strains 90 & 21"
 /db_xref="taxon:9031"
 /clone="pgm2n.pk008.l2"
 /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
 and Epiphyseal Growth Plate cDNA library (pgm2n)"
 /sex="Male and Female"
 /tissue_type="Breast muscle, leg muscle and epiphyseal
 growth plate"
 /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
 ,11 weeks);growth plate(1d,7d,14d post-hatch)"
 /lab_host="E. coli EMDH10B"
 /note="Vector: pCMVSPORT6; Library made from equivalent
 pools of total RNA isolated from each tissue (embryonic
 muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
 plate 33.3% of the final RNA pool). Single pass sequencing
 from 5'-end"

BASE COUNT 90 a 183 c 197 g 70 t 6 others

ORIGIN

Query Match 68.6%; Score 195.4; DB 10; Length 546;
 Best Local Similarity 80.7%; Pred. No. 5.7e-50;
 Matches 226; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 1 atcgagagtctacagcgtgacccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
 Db 217 ATGCAGAGCCTGCGCCGCCGCGCTACCGCAACGCTGCGAGGCGCTGTGG 276
 Qy 61 aggattataagaacggaggcctatgagggccatgaggggctgaacgtcacagcaaca 120
 Db 277 CGCATCTGCGTACCGAGGGCGTGTGAGGCCCATGCGGGGCTGAACATCACCCGCC 336
 Qy 121 ggcagaggcctgcccagcccttatttgcctgctacgaaaaagttaaaagacattg 180
 Db 337 GCGCGGGGCGCGCCCGCCCTCTACTTCGCTGCTACGAAAAAGTTAAAAAAGACGCTG 396
 Qy 181 agtgatgaattccacctgggggcaatagccattatcccaatgctgcgaggtgtgtg 240
 Db 397 AGCAGCTCATCCAGCGGGGGCAATAGCCATGTGCNNCGGTGCGCGGGTGTGTA 456
 Qy 241 gcaacattacttcattgatgcagccatgaacctgcggaag 280
 Db 457 GCAACATTGCTCCAGCAGCGCAGCGATGAACCCCTGCAGAAG 496

RESULT 15

BM412175 855 bp mRNA linear EST 14-AUG-2001
 LOCUS 602965377F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5120755 5',
 DEFINITION mRNA sequence.

ACCESSION BM412175

VERSION BM412175.1 GI:15173098

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 855)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1294 row: j column: 20
High quality sequence start: 7
High quality sequence stop: 809.
FEATURES
Location/Qualifiers
1..855
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5120755"
/clone_lib="NCI CGAP Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT73B-pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCTCTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 209 a 233 c 241 g 172 t
ORIGIN

Query Match 67.3%; Score 191.8; DB 10; Length 855;
Best Local Similarity 90.6%; Pred No. 8.4e-49;
Matches 260; Conservative 0; Mismatches 22; Indels 5; Gaps 5;
Qy 1 atgcagagtctacagccctgacccag-ctgccgcgtatcgcaatgtgtgaggccctctg 59
Db 86 ATGCAGAGCCTACAGCCTGACCCAGACCGCCGCTATCGGAACGTTGTGGAGGCTCTCTG 145
Qy 60 gaggtattataagaacggaggccctatgaggcccatgagggg-gctgaacgtcacagcaa 118
Db 146 GAGATCATGAGGAGAGAGGGCCTGTGGAGGCCCATCGCGGGTGTGTAACGTACAGCAA 205
Qy 119 caggcgagggcc-tgcccacgcccct-ttattttgcctgctacgaaagttaaaaaagac 176
Db 206 CAGGCGGGGGCGCTGCCACGCCCTCGTATTTGCCCTGCTACGAAAAGTTAAAAAGAC 265
Qy 177 a-ttgagtgtatgaatccaccctggggcaatagccatattgccaatggtcgccgggt 235
Db 266 AGTTGAGTGACGTAATCCACCAGGGGGCAATAGCCATATTGCCAATGTGCAGCCGGAT 325
Qy 236 gtgtggaacattacttcattatgcagccatgaacccctcggaagggc 282
Db 326 GTGTGGGACATTACTTCATGATGCAGCCATGAATCCAGCGGAAGTC 372

Search completed: September 28, 2002, 06:22:49
Job time: 9190 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:58:26 ; Search time 5287.54 Seconds
(without alignments)
1567.252 Million cell updates/sec

Title: US-09-870-113-11

Perfect score: 396

Sequence: 1 atgcagagtgctacagcctga.....ttacctgtttaccactag 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vl.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.Other.*

33: em.htgo.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	280	70.7	1322	6	AX061229	Sequence
2	280	70.7	1448	9	AF327402	Homo sapi
3	280	70.7	1448	17	HS303077	Homo sapi
4	273	68.9	401	6	AX071440	Sequence
5	225	56.8	1889	9	AF327403	Homo sapi
6	225	56.8	1889	17	HS303078	Homo sapi
7	225	56.8	123160	9	AL353719	Human DNA
8	202.6	51.2	221062	2	AC096351	Rattus no
9	139.2	35.2	4112	10	AF288621	Homo sapi
10	128	32.3	1429	9	AF223466	Mus muscu
11	125.2	31.6	163205	2	AC099368	Rattus no
12	109.8	27.7	5869	10	AF361699	Mus muscu
13	109.8	27.7	191702	2	AC108878	Mus muscu
14	105	26.5	108765	2	AF216674	Homo sapi
15	104.8	26.5	483	6	AX306258	Sequence
16	103.4	26.1	858	9	BC015013	Homo sapi
17	103.4	26.1	3415	9	AY032628	Homo sapi
18	103.4	26.1	174445	9	AC051642	Homo sapi
19	103.4	26.1	182504	2	AC022597	Homo sapi
20	60.6	15.3	30891	2	AC100095	Mus muscu
21	50	12.6	1072	10	BC023172	Mus muscu
22	49.6	12.5	34305	3	CEW02912	Caenorhabdi
23	42.2	10.7	3037	3	AY060268	Drosophila
24	40.6	10.3	2878	3	AF217402	Drosophila
25	36.2	9.1	1243	3	OGU45997	Onchocerca
26	36.2	9.1	1281	3	OU45998	Onchocerca
27	36	9.1	111464	8	F10K1	Sequence
28	35.6	9.0	180397	2	AP003615	Oryza sat
29	35.2	8.9	90935	2	AP003917	Oryza sat
30	35.2	8.9	169672	2	AC103286	Rattus no
31	35	8.8	5747	9	AF108460	Homo sapi
32	35	8.8	6334	9	AF108461	Homo sapi
33	34.6	8.7	223538	2	AC087150	Mus muscu
34	34.4	8.7	1517	9	AF125531	Homo sapi
35	34.4	8.7	205879	2	AL669956	Mus muscu
36	34.2	8.6	28993	3	LMFL5808	Plasmodium
37	34	8.6	373	3	AF221826	Leishmani
38	34	8.6	657	8	AF121142	Gracillari
39	34	8.6	7242	8	MITTARRNG	T.aestivum
40	34	8.6	103787	9	AC093659	Homo sapi
41	34	8.6	152248	2	AC027522	Homo sapi
42	34	8.6	184034	2	AC024156	Homo sapi
43	33.8	8.5	1154	4	AF127029	Bos tauru
44	33.8	8.5	52639	2	AC107472	Rattus no
45	33.8	8.5	64790	2	AC099356	Rattus no

ALIGNMENTS

RESULT	1	AX061229	Sequence	76 from Patent WO0078953.	1322 bp	DNA	linear	PAT 22-JAN-2001
LOCUS	AX061229	Sequence	76 from Patent WO0078953.					
DEFINITION	AX061229	Sequence	76 from Patent WO0078953.					
ACCESSION	AX061229	Sequence	76 from Patent WO0078953.					
VERSION	AX061229.1	Sequence	76 from Patent WO0078953.					
KEYWORDS	AX061229.1	Sequence	76 from Patent WO0078953.					
SOURCE	AX061229.1	Sequence	76 from Patent WO0078953.					
ORGANISM	AX061229.1	Sequence	76 from Patent WO0078953.					
REFERENCE	AX061229.1	Sequence	76 from Patent WO0078953.					
AUTHORS	AX061229.1	Sequence	76 from Patent WO0078953.					
TITLE	AX061229.1	Sequence	76 from Patent WO0078953.					
JOURNAL	AX061229.1	Sequence	76 from Patent WO0078953.					
FEATURES	AX061229.1	Sequence	76 from Patent WO0078953.					
source	AX061229.1	Sequence	76 from Patent WO0078953.					
BASE COUNT	AX061229.1	Sequence	76 from Patent WO0078953.					

ORIGIN

Query Match 70.7%; Score 280; DB 6; Length 1322;
 Best Local Similarity 100.0%; Pred. No. 1.2e-81;
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgacccagctgcccgcctatgcgaatgtgttgagggccctctgg 60
 |||||
 Db 186 ATGCAGAGCTACAGCCTGACCCAGCTGCCGCTATGCAATGTGTGGAGGCCCTCTGG 245
 |||||

Qy 61 aggattataagaacgagggcctatggagggccatgaggggctgaacgtcacagcaaca 120
 |||||
 Db 246 AGGATTATAAGAACGAGGGGCTATGGAGGCCCATGAGGGGCTGAAGGTACAGCAACA 305
 |||||

Qy 121 ggcgagggcctgcccagccttatttgcctgctacgaaagttaaaaaaacacattg 180
 |||||
 Db 306 GCGCAGGGCCTGCCACGCCCTTTATTTGCTGCTACGAAAGTTAAAAAAGACATTG 365
 |||||

Qy 181 agtgatgtaattccaccctgggggcaatagccatattgccaatgtgcccgggtgtgtg 240
 |||||
 Db 366 AGTGATGTAATCCACCCCTGGGGCAATAGCCATATTGCCAATGTGCGCGCGGTGTGTG 425
 |||||

Qy 241 gcaacattacttcgatgcagccatgaacccctgcggaag 280
 |||||
 Db 426 GCAACATTACTTTCATGATGCAGCCATGAACCCCTGCGGAAG 465
 |||||

RESULT 2

AF327402 1448 bp mRNA linear PRI 02-MAY-2001
 LOCUS AF327402
 DEFINITION Homo sapiens putative mitochondrial solute carrier splice variant mRNA, complete cds, alternatively spliced, nuclear gene for mitochondrial product.

ACCESSION AF327402.1 GI:13926046

VERSION AF327402

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1448)

Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B.,

Schweyen, R., Larsson, C. and Suomalainen, A.

Characterization of a novel human putative mitochondrial

transporter homologous to the yeast mitochondrial RNA splicing

proteins 3 and 4

FEBS Lett. 494 (1-2), 79-84 (2001)

21195335

11297739

2 (bases 1 to 1448)

Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B., Schweyen, R.,

Larsson, C. and Suomalainen, A.

Direct Submission

Submitted (08-DEC-2000) Human Molecular Genetics, National Public

Health Institute, Mannerheimintie 166, Helsinki 00300, Finland

Location/Qualifiers

1. 1448

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="10"

/map="10q24"

30. 1124

/note="similar to mitochondrial RNA splicing protein 3/4;

alternatively spliced"

/codon_start=1

/product="putative mitochondrial solute carrier splice

variant"

/protein_id="AAK49519.1"

/db_xref="GI:13926047"

/translation="MELEGRGAGVAGGPAAGPGRSPGESALLDGLORVGRGAGG

EAGACRPVQRQDSGDYDALPAGATVTHWVAGVAGLHCVMYPCDVYTRMQS

LQPDPAARYNLEALWRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYERLUKTKLS

DVTHPGNSHIANGAAGCVATLLHDAAMPNPAEVVVKORMQMYNSPYHRVTDVRAVQON
 EGAGAFYRSYTTQLTNNVPFQAIHFNTYFLQEHFNPPRRYNPSSHVLSGACAGAVAA
 AATTPLDVKCKTLNLTNESHITGHITGMASAFRTVYQGVGTAYFRGQARVIY
 QIPSTAIANSVVEFFKYLITKROEWRAG"

BASE COUNT 322 a 408 c 435 g 283 t
 ORIGIN

Query Match 70.7%; Score 280; DB 9; Length 1448;
 Best Local Similarity 100.0%; Pred. No. 1.3e-81;
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgacccagctgcccgcctatgcgaatgtgttgagggccctctgg 60
 |||||
 Db 327 ATGCAGAGCTACAGCCTGACCCAGCTGCCGCTATGCAATGTGTGGAGGCCCTCTGG 386
 |||||

Qy 61 aggattataagaacgagggcctatggagggccatgaggggctgaacgtcacagcaaca 120
 |||||
 Db 387 AGGATTATAAGAACGAGGGGCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 446
 |||||

Qy 121 ggcgagggcctgcccagccttatttgcctgctacgaaagttaaaaaaacacattg 180
 |||||
 Db 447 GGCAGGAGGCTGCCACGCCCTTTATTTGCTGCTACGAAAGTTAAAAAAGACATTG 506
 |||||

Qy 181 agtgatgtaattccaccctgggggcaatagccatattgccaatgtgcccgggtgtgtg 240
 |||||
 Db 507 AGTGATGTAATCCACCCCTGGGGCAATAGCCATATTGCCAATGTGCGCGGTGTGTG 566
 |||||

Qy 241 gcaacattacttcgatgcagccatgaacccctgcggaag 280
 |||||
 Db 567 GCAACATTACTTTCATGATGCAGCCATGAACCCCTGCGGAAG 606
 |||||

RESULT 3

HSA303077

ID HSA303077 standard; RNA; HUM; 1448 BP.

XX

AC AJ303077;

XX

SV AJ303077.1

XX

DT 02-FEB-2001 (Rel. 56, Created)

DT 26-APR-2001 (Rel. 67, Last updated, Version 2)

XX

DE Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMRS3/4

gene), 1448 bp cDNA splice variant

DE

XX

KW HMRS3/4 gene; mitochondrial RNA splicing protein 3/4.

XX

OS Homo sapiens (human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX

XX [1]

RP 1-1448

RA Nikali K.;

RT ;

RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.

RL Nikali K., Human Molecular Genetics, National Public Health Institute,

Mannerheimintie 166, Helsinki, FI-00300, FINLAND.

XX

XX [2]

RA Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,

Larsson C., Suomalainen A.;

RT "Characterization of a novel human putative mitochondrial transporter

homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";

RL FEBS Lett. 494:79-84 (2001).

XX

Key Location/Qualifiers

1. 1448

/db_xref="taxon:9606"

/organism="Homo sapiens"

FT

```

CDS
30. .1124
/ gene="HMR3/4"
/ product="mitochondrial RNA splicing protein 3/4"
/ function="putative mitochondrial solute carrier"
/ protein_id="CAC27996.1"
/ translation="MELGRGAGGAVAGGPAAGPGRSPGESALLDGLRGVGRGAGGGE
AGACPPVRQDPDSDPDEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQ
PDPAARYRNLEALNRIITREGLRPMRLNVTATGAGPAHALYFAECKLKLTSVDVI
HPGNSHTANGAAGCAVATLLHDAANPAEVVKRQMYNSPVHRVTDCAVRVWNEGAG
AFYRSYTTQLMNVFOAHEFMEFLOEHEFNPPRRYNSPSHVLSCAGAVAAATTP
LDVCKTLNTOESLALNSHITGMSAFRTVYQVGGVATYFRGVQARVIYQIPSTA
TANSYEFKYLITRQEWRAKG"
SQ Sequence 1448 BP; 322 A; 408 C; 435 G; 283 T; 0 other;

Query Match 70.7%; Score 280; DB 17; Length 1448;
Best Local Similarity 100.0%; Pred. No. 1.3e-81;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagtctacagctgaccacgctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
DB 327 ATGCAGAGTCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 386

QY 61 agattataagaacgagggccctatgagggccatgaggggctgaacgtcacagcaaca 120
|||||
DB 387 AGGATTATAGAACGAGGGCCCTATGGAGGCCCATGAGGGGGCTGAACGTACAGCAACA 446

QY 121 ggcgagggcctgccacgccctttatttgcctgctacgaaagttaaaaagacattg 180
|||||
DB 447 GGCAGGGGCCCTGCCACGCCCTTTATTTGGCTGCTACGAAAGTTAAAAAGACATTG 506

QY 181 agtgatgtaatccaccctgggggcaatagccattgccaatggtgcggcggtgtgtg 240
|||||
DB 507 AGTGATGTAATCCACCCTGGGGCAATACCCATATTGCCAATGGTGCGCCGGGTGTG 566

QY 241 gaacattacttcattgacgacccatgaacctgcggaag 280
|||||
DB 567 GCAACATTACTTCATGATGATGACCCATGAACCCCTGCGGAAG 606

RESULT 4
AX071440
LOCUS AX071440 401 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 1912 from Patent WO0102568.
ACCESSION AX071440
VERSION AX071440.1 GI:12581791
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 401)
AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
Lanson, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M.,
Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 1912 11-JAN-2001;
CHIRON CORPORATION (US); HYSEQ, INC. (US)
FEATURES
Location/Qualifiers
source 1..401
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1..401
/note="n = A,T,C or G"
BASE COUNT 100 a 104 c 120 g 76 t 1 others
ORIGIN

Query Match 68.9%; Score 273; DB 6; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.1e-79;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

```

QY 8 gtctacagctgacccagctgcccgtatcgcaatgtgttgaggccctctgaggatta 67
|||||
DB 9 GTCTACAGCTGAOCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTCGAGGATTA 68

QY 68 taagaacgagggccctatgagggccatgaggggctgaacgtcacagcaacagcgag 127
|||||
DB 69 TAAGAACGAGGGCCCTATGAGGCCCATGAGGGGGCTGAACGTACAGCAACAGCGCAG 128

QY 128 ggcctgccacgccctttatttgcctgctacgaaagttaaaaagacattgagtgtg 187
|||||
DB 129 GGCCTGCCACGCCCTTTATTTGGCTGCTACGAAAGTTAAAAAGACATTGAGTGATG 188

QY 188 taatccaccctgggggcaatagccattgccaatggtgcggcggtgtgtggcaaat 247
|||||
DB 189 TAATCCACCCTGGGGCAATAGCCATATTGCCAATGGTGCGCCGGGTGTGTGGCAACAT 248

QY 248 tacttcattgacgacccatgaacctgcggaag 280
|||||
DB 249 TACTTCATGATGACGCATGAACCCCTGCGGAAG 281

RESULT 5
AF327403
LOCUS AF327403 1889 bp mRNA linear PRI 02-MAY-2001
DEFINITION Homo sapiens putative mitochondrial solute carrier splice variant
mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.
ACCESSION AF327403
VERSION AF327403.1 GI:13926049
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1889)
AUTHORS Li, F.Y., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B.,
Schweyen, R., Larsson, C. and Suomalainen, A.
TITLE Characterization of a novel human putative mitochondrial
transporter homologous to the yeast mitochondrial RNA splicing
proteins 3 and 4
JOURNAL FEBS Lett. 494 (1-2), 79-84 (2001)
MEDLINE 21195335
PUBMED 11297739
REFERENCE 2 (bases 1 to 1889)
AUTHORS Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B., Schweyen, R.,
Larsson, C. and Suomalainen, A.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-2000) Human Molecular Genetics, National Public
Health Institute, Mannerheimintie 166, Helsinki 00300, Finland
FEATURES
Location/Qualifiers
source 1..1889
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q24"
1032..1565
/note="similar to mitochondrial RNA splicing protein 3/4;
alternatively spliced"
/codon_start=1
/product="putative mitochondrial solute carrier splice
variant"
/protein_id="AAK49520.1"
/db_xref="GI:13926050"
/translation="MNPAAVVKRQMYNSPVHRVTDCAVRVWNEGAGAFYRSYTTQ
LTMNVFOAHEFMEFLOEHEFNPPRRYNSPSHVLSCAGAVAAATTPIDVCKTL
NTQESLALNSHITGMSAFRTVYQVGGVATYFRGVQARVIYQIPSTAIAVSVLE
FEKYLITRQEWRAKG"
BASE COUNT 404 a 508 c 544 g 433 t
ORIGIN

Query Match 56.8%; Score 225; DB 9; Length 1889;


```

```

Best Local Similarity 100.0%; Pred. No. 2.5e-63;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctacagcctgaccagctgcccgcctatcgaatgtgttgaggccctctgg 60
DB 306 ATGCAGAGCTACAGCCTGACCCAGCTGCCCGCTATCGCAATGTGTGGAGGCCCTCTGG 365
QY 61 aggattataagaacagagggcctatggagcccatgaggggctgaacgtcacagcaaca 120
DB 366 AGGATTATAAGAACAGGAGGCGCTATGGAGGCCCATGAGGGGCGCTGAACGTACACGACACA 425
QY 121 ggcgagggcctgcccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 180
DB 426 GCGCAGGGCCTGCCACGCCCTTTATTTCCTGCTGCTACGAAAGTTAAAAAGACATTG 485
QY 181 agtgatgtaataccacctgggggcaatgacgaattgccaatggt 225
DB 486 AGTGATGTAATCCACCTGGGGCAATAGCCATATTGCCAATGGT 530

RESULT 6
HSA303078
ID HSA303078 standard; RNA; HUM; 1889 BP.
AC AJ303078;
SV AJ303078.1
XX
XX
XX
XX 02-FEB-2001 (Rel. 66, Created)
XX 26-APR-2001 (Rel. 67, Last updated, Version 2)
XX
XX Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMR3/4
DE gene), 1889 bp cDNA splice variant
XX
XX HMR3/4 gene; mitochondrial RNA splicing, protein 3/4.
XX
XX Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
XX [1]
RN 1-1889
RA Nikali K.;
RT ;
RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.
RL Nikali K., Human Molecular Genetics, National Public Health Institute,
RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.
XX
XX [3]
RA Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
RA Larsson C., Suomalainen A.;
RT "Characterization of a novel human putative mitochondrial transporter
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";
RL FEBS Lett. 494:79-84(2001).
XX
XX Key Location/Qualifiers
FH
FT 1..1889
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT CDS 1032..1565
FT /gene="HMR3/4"
FT /product="mitochondrial RNA splicing protein 3/4"
FT /function="putative mitochondrial solute carrier"
FT /protein_id="CAC27997.1"
FT /translation="MNFAEVVKQMNYNSPHRVTDICRAVWQNEGAGAFYSYTTQL
FT TMNVPFQAIHMTYEFQLQEHNPORRYNPSHVLGACAGAAAAATTPLDVCXTLLNT
FT QESLALNSHTGHTGMASAPRTYVQGGVTAYFRGVQARVIYQIPSTAIWSVYEFK
FT YLITKQEWNRAGR"
XX
XX Sequence 1889 BP; 404 A; 508 C; 544 G; 433 T; 0 other;

```

```

Query Match 56.8%; Score 225; DB 17; Length 1889;
Best Local Similarity 100.0%; Pred. No. 2.5e-63;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctacagcctgaccagctgcccgcctatcgaatgtgttgaggccctctgg 60
DB 306 ATGCAGAGCTACAGCCTGACCCAGCTGCCCGCTATCGCAATGTGTGGAGGCCCTCTGG 365
QY 61 aggattataagaacagagggcctatggagcccatgaggggctgaacgtcacagcaaca 120
DB 366 AGGATTATAAGAACAGGAGGCGCTATGGAGGCCCATGAGGGGCGCTGAACGTACACGACACA 425
QY 121 ggcgagggcctgcccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 180
DB 426 GCGCAGGGCCTGCCACGCCCTTTATTTCCTGCTGCTACGAAAGTTAAAAAGACATTG 485
QY 181 agtgatgtaataccacctgggggcaatgacgaattgccaatggt 225
DB 486 AGTGATGTAATCCACCTGGGGCAATAGCCATATTGCCAATGGT 530

RESULT 7
AL353719/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete
sequence.
ACCESSION AL353719 AC007643
VERSION AL353719.10 GI:15787725
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 123160)
Direct Submission
Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 26, 2001 this sequence version replaced gi:14280413.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em; EMBL; SW;
SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-85A1 is from the library RP11-11.1 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-85A1. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP11-85A1 is at 1 in this sequence. The
true left end of clone RP11-483F11 is at 123061 in this sequence.
The true right end of clone RP11-129J12 is at 51589 in this
sequence.

```

```

FEATURES
  source      Location/Qualifiers
1. .123160
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="10"
   /clone="RP11-85A1"
   /clone_lib="RPC1-11.1"
   complement(7066..7118)
   /note="Single clone region. Sequence from reads from a
   short insert library derived from a single pUC clone.
   Restriction digest data confirm the assembly."
52162
52167
  misc_feature
  unsure
  unsure
  BASE COUNT      33221 a 28084 c 27196 g 34659 t
  ORIGIN

Query Match      56.8%; Score 225; DB 9; Length 123160;
Best Local Similarity 100.0%; Pred. No. 5.1e-63;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagctgaccagctccgcgtatcgaatgtgttgagccctctgg 60
|||||
Db 85540 ATGCAGAGTCTACAGCTGACCCAGCTCCCTCTATCGCAATGTTGTGGAGCCCTCTGG 85481
|||||

Qy 61 aggtattataagaacgagggcctatggggccatgagggggtgaacgtcacagcaaca 120
|||||
Db 85480 AGGATTATAAGAACGAGGCGCTATGGAGGCCATGAGGGGGCTGACGTCACACCAACA 85421
|||||

Qy 121 ggcgcagggcctgcccacgcccctttatttgcctgctacgaaagttaaaaagacattg 180
|||||
Db 85420 GCGCAGAGCGCTGCGCACGCGCCCTTTATTTCCTGCTACGAAAGTTAAAAAGACATTG 85361
|||||

Qy 161 agtgaatgaatccacctgggggcaatagccatatgtgcaatgggt 225
|||||
Db 85360 ACTGATGTAATCCACCCTGGGGGCAATAGCCCATATFTGCCAATGGT 85316
|||||

RESULT      8
AC096351/c
LOCUS
DEFINITION
Rattus norvegicus chromosome R1 clone CH230-24M6, WORKING DRAFT
SEQUENCE, 33 unordered pieces.
ACCESSION
AC096351.2 GI:17944054
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Norway rat.
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.B., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle-McDoll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

```

```

Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,A., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.I.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley,K.C.
Direct Submission
Submitted (17-Sep-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 32851: contig of 32851 bp in length
* 32852 32951: gap of unknown length
* 32952 48984: contig of 16033 bp in length
* 48985 49084: gap of unknown length
* 49085 63815: contig of 14731 bp in length
* 63816 63915: gap of unknown length
* 63916 76005: contig of 12090 bp in length
* 76006 76105: gap of unknown length
* 76106 88841: contig of 12736 bp in length
* 88842 88941: gap of unknown length
* 88942 100591: contig of 11650 bp in length
* 100592 100691: gap of unknown length
* 100692 111702: contig of 11011 bp in length
* 111703 111802: gap of unknown length
* 111803 122664: contig of 10862 bp in length
* 122665 122764: gap of unknown length
* 122765 129711: contig of 6947 bp in length
* 129712 129811: gap of unknown length

```



```

RESULT 10
AF223466
LOCUS      Homo sapiens HT015 protein (HT015) mRNA, complete cds.
DEFINITION
ACCESSION AF223466
VERSION    AF223466.1 GI:7578782
KEYWORDS   human.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 1429)
            Song, H., Gao, G., Peng, Y., Ren, S., Chen, Z. and Han, Z.
            A novel gene expressed in human hypothalamus
            Unpublished
REFERENCE  2 (bases 1 to 1429)
            Song, H., Gao, G., Peng, Y., Ren, S., Chen, Z. and Han, Z.
            Direct Submission
            Submitted (12-JAN-2000) Chinese National Human Genome Center at
            Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, China
FEATURES   Location/Qualifiers
            1..1429
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /tissue_type="hypothalamus"
            1..1429
                /gene="HT015"
                /gene_id="AAF64141.1"
                /protein_id="AAF64141.1"
                /db_xref="GI:7578783"
                /translation="MELRSGVSGQAVARMGDSRGGGKDGATGSEYENLPTSS
                VSTHMTAGMAGIIEHVMYVDSVKTRMOSLSDPKAQYTSIYGALKKINRTGFWR
                PLRGVNYIMQAGPAHAMYFAICYENMKRTLNDVPHQGNHSLANGIAGSMATLLHDVA
                MNPAYVYKQNGNSGHRSAICRTVWRTGELGAFYRSYVTTOLTWNIPQSLHFT
                YEFQEOVNPRTYVPOSHIISGGLAGALAAATPLDVCCKTLTNTQENVALSIANTIS
                GRLSGMANPRTVYOLNGLPATSKASRRVSTRCPPPFPFLGLSSSSSTSPSPASWKI
                ELHFKGRDHRIFS"
            326 a 426 c 402 g 275 t
BASE COUNT
ORIGIN
Query Match      32.3%; Score 128; DB 9; Length 1429;
Best Local Similarity 66.1%; Pred. No. 4.le-31;
Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
Qy 1 atcgagagtctacagcctgacccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
Db 261 ATGCAGAGTTTGAGTCCAGATCCCAAGCCAGTACACAAGTATCTACGGAGCCCTCAAG 320
Qy 61 aggattataaagcaggaggcctatggaggccatgagggcgctgaacgtcacacaca 120
Db 321 AAAATCATGCGGACCAAGAGGCTTCTGGAGGCCCTTTCGAGAGCGCTCAACGTATGATCATG 380
Qy 121 ggcgcaggcctgcccgcgccttatttgcctgctacgaaagttaaaaaagacattg 180
Db 381 GGTGCGAGGCGCGCCCATGACCATGATATTTTTCCTGCTATGAAACATGAAAGGACTTTA 440
Qy 181 agtgaatgaatccaccctggggccaatagccatattgccaatggcgccgggtgtgtg 240
Db 441 AATGACGTTTTCCACCACCAAGAAACAGCCACCCTAGCCAGGGATAGTGGGAGTATG 500
Qy 241 gcaacattacttcattgatgcagcgaatgaacccctgggaag 280
Db 501 GCCACCTGCTCCACGATCGCGTATGATGAATCCAGCAGAAG 540

LOCUS      163205 bp      DNA      linear      HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-84G11, *** SEQUENCING IN PROGRESS
            ***, 71 unordered pieces.
ACCESSION AC099368
VERSION    AC099368.2 GI:17973175
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 163205)
            Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
            Albrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J.,
            Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J.,
            Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
            Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
            Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
            Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
            Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
            Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
            Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
            Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
            Elhaj, C., Escotto, M., Falls, T., Ferraguto, B., Flagg, N., Ford, J.,
            Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
            Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
            Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
            Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C.,
            Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
            Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
            Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
            Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
            Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
            Loulseg, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
            Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
            Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,
            Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
            Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
            Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,
            Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
            Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
            Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,
            Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
            Sisson, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H.,
            Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
            Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
            Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,
            Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
            Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
            Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
            Weinstock, G. and Gibbs, R.
            Direct Submission
            Unpublished
            2 (bases 1 to 163205)
            Worley, K.C.
            Direct Submission
            Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On Dec 20, 2001 this sequence version replaced gi:16901901.
            ----- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ----- Project Information
            Center project name: GH1J
            Center clone name: CH230-84G11
            ----- Summary Statistics
            Assembly program: Phrap; version 0.990329First call to
            findPhrapList
            Consensus quality: 127058 bases at least Q40
            Consensus quality: 138264 bases at least Q30

```

Consensus quality: 146548 bases at least Q20
Estimated insert size: 120308; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 8509: contig of 8509 bp in length
* 8609: gap of unknown length
* 8610: contig of 4931 bp in length
* 13541: 13640: gap of unknown length
* 13641: 19516: contig of 5876 bp in length
* 19517: 19616: gap of unknown length
* 19617: 24990: contig of 5374 bp in length
* 24991: 25090: gap of unknown length
* 25091: 29401: contig of 4311 bp in length
* 29402: 29501: gap of unknown length
* 29502: 33153: contig of 3552 bp in length
* 33154: 38439: contig of 5286 bp in length
* 38440: 38539: gap of unknown length
* 38540: 41189: contig of 2644 bp in length
* 41184: 41283: gap of unknown length
* 44936: 45036: contig of 3653 bp in length
* 45037: 47637: contig of 2601 bp in length
* 47638: 47737: gap of unknown length
* 47738: 50166: contig of 2429 bp in length
* 50167: 50266: gap of unknown length
* 50267: 54170: contig of 3904 bp in length
* 54171: 54270: gap of unknown length
* 54271: 57672: contig of 3402 bp in length
* 57672: 57772: gap of unknown length
* 57773: 61162: contig of 3390 bp in length
* 61163: 61262: gap of unknown length
* 61263: 63890: contig of 2628 bp in length
* 63891: 63990: gap of unknown length
* 63991: 66900: contig of 2910 bp in length
* 66901: 67000: gap of unknown length
* 67001: 69717: contig of 2717 bp in length
* 69718: 69817: gap of unknown length
* 69818: 72921: contig of 3104 bp in length
* 72922: 73021: gap of unknown length
* 73022: 75145: contig of 2125 bp in length
* 75147: 75246: gap of unknown length
* 75247: 79038: contig of 3792 bp in length
* 79039: 79138: gap of unknown length
* 79139: 80340: contig of 1202 bp in length
* 80341: 80440: gap of unknown length
* 80441: 82366: contig of 1926 bp in length
* 82367: 82466: gap of unknown length
* 82467: 84184: contig of 1718 bp in length
* 84185: 84284: gap of unknown length
* 84285: 86001: contig of 1717 bp in length
* 86002: 86101: gap of unknown length
* 86102: 87695: contig of 1594 bp in length
* 87696: 87795: gap of unknown length
* 87796: 89287: contig of 1492 bp in length
* 89288: 89387: gap of unknown length
* 89388: 91950: contig of 2563 bp in length
* 91951: 92050: gap of unknown length
* 92051: 94575: contig of 2525 bp in length
* 94576: 94673: gap of unknown length
* 94674: 96581: contig of 1906 bp in length
* 96582: 96681: gap of unknown length

96682 98235: contig of 1554 bp in length
* 98236: gap of unknown length
* 98336 98814: contig of 1479 bp in length
* 98914: gap of unknown length
* 99915 100964: contig of 1050 bp in length
* 100965 101064: gap of unknown length
* 101065 103074: contig of 2010 bp in length
* 103075 103174: gap of unknown length
* 103175 104311: contig of 1137 bp in length
* 104312 104411: gap of unknown length
* 104412 106834: contig of 2423 bp in length
* 106835 106934: gap of unknown length
* 106935 108922: contig of 1988 bp in length
* 108923 109022: gap of unknown length
* 109023 110868: contig of 1846 bp in length
* 110869 110968: gap of unknown length
* 110969 113056: contig of 2088 bp in length
* 113057 113156: gap of unknown length
* 113157 114296: contig of 1140 bp in length
* 114297 114396: gap of unknown length
* 114397 116312: contig of 1916 bp in length
* 116313 116412: gap of unknown length
* 116413 117492: contig of 1080 bp in length
* 117493 117592: gap of unknown length
* 117593 119552: contig of 1960 bp in length
* 119553 119652: gap of unknown length
* 119653 120658: contig of 1006 bp in length
* 120659 120758: gap of unknown length
* 120759 121895: contig of 1137 bp in length
* 121896 121995: gap of unknown length
* 121996 123272: contig of 1277 bp in length
* 123273 123372: gap of unknown length
* 123373 125310: contig of 1938 bp in length
* 125311 125410: gap of unknown length
* 125411 126891: contig of 1481 bp in length
* 126892 126991: gap of unknown length
* 126992 128690: contig of 1699 bp in length
* 128691 128790: gap of unknown length
* 128791 130827: contig of 2037 bp in length
* 130828 130927: gap of unknown length
* 130928 132501: contig of 1574 bp in length
* 132502 132601: gap of unknown length
* 132602 134071: contig of 1470 bp in length
* 134072 134171: gap of unknown length
* 134172 135838: contig of 1667 bp in length
* 135839 135938: gap of unknown length
* 135939 137589: contig of 1651 bp in length
* 137590 137689: gap of unknown length
* 137690 139043: contig of 1354 bp in length
* 139044 139143: gap of unknown length
* 139144 140777: contig of 1634 bp in length
* 140778 140877: gap of unknown length
* 140878 142526: contig of 1649 bp in length
* 142527 142626: gap of unknown length

Query Match 31.6%; Score 125.2; DB 2; Length 163205;
Best Local Similarity 87.4%; Pred. No. 7.9e-30;
Matches 194; Conservative 0; Mismatches 23; Indels 5; Gaps 5;

Qy 7 agtctaacgctgaccag-ctgccgcctatcgcaatgtgtgga-ggcccctctggagga 64
Db 13720 ATTTTACAGCCTGATCCAGCGCGCCCTTATCGAAAAAGTTGTGGAGGCTTTTGGAGAA 13779
Qy 65 ttataagaacggaggcctatggaggcccatga-gggggcgtgaacgtcacagcaacagc 123
Db 13780 TTATAAGACAGAGCGCCCTGTGGAGGCCCATGAGGGGGTGTGACCTGACCTGCCTGCG 13839
Qy 124 gcaggcctgcccacgccccttatttgcctgcctacgaaaaagttaaaaaagacattagat 183
Db 13840 CGGGGGCTGCCACCGCCCTCTATTTTG-CTGCTACGAAAAAG-TAAAAAGACATTGAGT 13897
Qy 184 gatgtaatccaccctgggggcaatagccatatgtccaatggt 225
|||||

```

Db 13898 GATGTAATCCACCCGGGGCAATAGCCATATTGCCAATGGT 13939
RESULT 12
AF361699          5869 bp      mRNA      linear      ROD 07-NOV-2001
LOCUS              Mus musculus mitochondrial carrier-like protein mRNA, complete cds;
DEFINITION          nuclear gene for mitochondrial product.
ACCESSION          AF361699
VERSION            AF361699.1  GI:16755527
KEYWORDS            house mouse.
SOURCE              Mus musculus
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS             Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and
                     She,J.
TITLE               A novel mouse mitochondrial carrier protein gene is up-regulated
                     from young to adult NOD mice
JOURNAL             Unpublished
REFERENCE            2 (bases 1 to 5869)
AUTHORS             Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and
                     She,J.
TITLE               Direct Submission
JOURNAL             Submitted (15-MAR-2001) Department of Pathology, Immunology and
                     Laboratory Sciences, University of Florida, 1600 SW Archer Road,
                     Room D6-15, Gainesville, FL 32610, USA
FEATURES             Location/Qualifiers
                     source            1..5869
                                     /organism="Mus musculus"
                                     /strain="C57BL/6J"
                                     /db_xref="taxon:10090"
                                     /chromosome="14"
                                     /map="between D14Mit124 and D14Mit236"
                                     116..664
                                     /note="MCLP"
                                     /codon_start=1
                                     /evidence=not_experimental
                                     /product="mitochondrial carrier-like protein"
                                     /protein_id="AAI27990.1"
                                     /db_xref="GI:16755528"
                                     /translation="MELPRGVGNQAGRRMDGCRGGCGSKDAGSEYENLPTSAS
VSTHMTAGAMGAILHSIMYPDVSKTMQSLNPDPKARYTSIYGALKRIMHTGFWR
PLRGLNMVMGAGPAHAMYFYACENMRKTLNDVFSHQGNHSLANGILKAFVWSWEALL
SGASSGPSNLHPROTENSRVT"
BASE COUNT          1387 a 1427 c 1390 g 1665 t
ORIGIN

Query Match          27.7%; Score 109.8; DB 10; Length 5869;
Best Local Similarity 68.0%; Pred. No. 5.9e-25;
Matches 153; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 atgcagagtctacagctgaccagctccgcgctatcgaatgtgttgaggccctctgg 60
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 332 ATGCAGAGTTTGAATCCAGATCCCAAGCCCGGTATACAGCATCTATGCGCCCTCAAG 391
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 aggattataagaagaggggccctatggagcccatgaggggctgaagcgtcacagcaaca 120
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 392 AGGATCATGCACATGAGAGGCTTCTGGAGGCCCTCGCGGGCCCTGAAGCTGATGATG 451
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 ggcgcaggccctgcccacgccccttatttgcctgctcagaaaagttaaaaagacattg 180
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 452 GGTGCAGGGCCCGGCACGCCCATGTATTTGCTCTCTATGAACAACATGAAAGGACTTTA 511
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 agtgatgtaatccacccttggggggaatggccatattgccaattg 225
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 512 AATGACGTTTTTCAGCCACCAAGAAACAGCCATCTAGCTAATGTT 556
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
AC108878

```

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AC108878 191702 bp DNA linear HTG 06-FEB-2002
Mus musculus clone RP23-256H24 strain C57BL6/J, *** SEQUENCING IN
PROGRESS ***, 17 unordered pieces.
AC108878
AC108878.2 GI:19542954
HTG: HTGS_PHASE1.
Mus musculus
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 191702)
Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J.,
Fusina,M., Halder,A., Keller,A., Perera,A., Shim,C., Thomas,E.,
Zencheck,W., Xi,C., Juels,P. and Kucheralapati,R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 191702)
Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J.,
Fusina,M., Halder,A., Keller,A., Perera,A., Shim,C., Thomas,E.,
Zencheck,W., Xi,C., Juels,P. and Kucheralapati,R.
Direct Submission
Submitted (01-FEB-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
On Feb 6, 2002 this sequence version replaced gi:18464039.
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site:
http://wchanning.bwh.harvard.edu:9088/hpcggg/jsp/hpcggg/Sequence/mous
e.html
Contact: gnktm@capecod.bwh.harvard.edu
-----Summary Statistics
Center project name: AKY
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 100%
*Consensus quality: 187211 at least Q20
*Consensus quality: 185727 at least Q30
*Consensus quality: 183181 at least Q40
*Estimated insert size: agarose-PP - N/A
**Estimated insert size: 191382 - sum-of-contigs
Quality coverage: agarose-PP - N/A
Quality coverage: 5.6 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 48676: contig of 48676 bp in length
* 48677 48696: gap of unknown length
* 48697 89065: contig of 40369 bp in length
* 89066 89085: gap of unknown length
* 89086 104994: contig of 15909 bp in length
* 104995 105014: gap of unknown length
* 105015 119612: contig of 14598 bp in length
* 119613 119632: gap of unknown length
* 119633 135969: contig of 16337 bp in length
* 135970 135989: gap of unknown length
* 135990 151803: contig of 15814 bp in length
* 151804 151823: gap of unknown length
* 151824 160052: contig of 8229 bp in length
* 160053 160072: gap of unknown length
* 160073 169021: contig of 8949 bp in length
* 169022 169041: gap of unknown length
* 169042 175745: contig of 6704 bp in length
* 175746 175765: gap of unknown length
* 175766 180321: contig of 4556 bp in length
* 180322 180341: gap of unknown length
* 180342 182588: contig of 2247 bp in length

182589 182608: gap of unknown length
182609 183975: contig of 1367 bp in length
183976 183995: gap of unknown length
183996 186393: contig of 2398 bp in length
186394 186413: gap of unknown length
186414 186416: contig of 2003 bp in length
186417 188436: gap of unknown length
188437 189584: contig of 1148 bp in length
189585 189604: gap of unknown length
189605 190648: contig of 1044 bp in length
190649 190668: gap of unknown length
190669 191702: contig of 1034 bp in length.
Location/Qualifiers
1..191702
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone="RP23-256H24"
/sex="male"
BASE COUNT 53763 a 43806 c 42665 g 51108 t 360 others
ORIGIN
Query Match 27.7%; Score 109.8; DB 2; Length 191702;
Best Local Similarity 68.0%; Pred. No. 1.1e-24;
Matches 153; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
Qy 1 atgcagagctacgctgaccgctgcccgcctatcgcaatgtgtggagccctctgg 60
Db 106283 ATGCAGAGCTTTGAATCCAGATCCCAAGCCGGTATCAAGCATCTATGGCGCCCTCAAG 106342
Qy 61 aggtattataagacgaggccctatgagccctatgagccctatgagggcctgaaacgctcacagcaaca 120
Db 106343 AGGATCATGCACATGAAGCTTCGAGGCCCCCTCGGGCCCTGAACGTGATGATGATG 106402
Qy 121 ggcagggcctgcccgccttatttgcctgctacgaaagttaaagacattg 180
Db 106403 GGTGACGGCCCGGCACGCCATCTATTTGCTGCTATGAACATGATAAGGACTTTA 106462
Qy 181 agtgaatgaatccacctgggggcaatagccatattgccaatggt 225
Db 106463 AATGACGTTTTACGCCCAACCAAGGAACAGCCATCTAGCTAATGTT 106507
RESULT 14
AF216674
LOCUS
DEFINITION Homo sapiens chromosome 8 clone RPI-158P13 map 8p, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
ACCESSION AF216674
VERSION AF216674.3 GI:14280182
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108765)
AUTHORS Schilhabel,M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Siddiqui,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,
Rosenthal,A. and Platzer,M.
Chromosome 8 genomic sequence
TITLE Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 108765)
AUTHORS Polley,A., Wen,G., Baumgart,C., Dette,M., Jahn,N., Schilhabel,M.,
Menzel,U. and Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-1999) Genome Analysis, Institute of Molecular
Biotechnology, Reutenbergstrasse 11, Jena 07745, Germany
COMMENT On Jun 2, 2001 this sequence version replaced gi:8151796.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB

Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H241
Center clone name: RPI-158P13
----- Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 87772 bases at least Q40
Consensus quality: 95159 bases at least Q30
Consensus quality: 100612 bases at least Q20
Quality coverage: 3,95 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1101: contig of 1101 bp in length
* 1201: gap of unknown length
* 21139: contig of 19938 bp in length
* 21140 21239: gap of unknown length
* 21240 36201: contig of 14962 bp in length
* 36202 36301: gap of unknown length
* 36302 47101: contig of 10800 bp in length
* 47102 47201: gap of unknown length
* 47202 56806: contig of 9605 bp in length
* 56807 56906: gap of unknown length
* 56907 65083: contig of 8177 bp in length
* 65084 65183: gap of unknown length
* 65184 73521: contig of 8338 bp in length
* 73522 73622 73621: gap of unknown length
* 81324 81423: gap of unknown length
* 81424 86503: contig of 5080 bp in length
* 86504 86603: gap of unknown length
* 86604 89473: contig of 2870 bp in length
* 89474 89573: gap of unknown length
* 89574 92369: contig of 2796 bp in length
* 92370 92470 92469: gap of unknown length
* 92471 94334: contig of 1965 bp in length
* 94335 96676: contig of 2142 bp in length
* 96677 96776: gap of unknown length
* 96777 98217: contig of 1441 bp in length
* 98218 98317: gap of unknown length
* 98318 101563: contig of 3246 bp in length
* 101564 101664: gap of unknown length
* 101664 108765: contig of 7102 bp in length.
Location/Qualifiers
1..108765
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p"
/clone="RPI-158P13"
BASE COUNT 28954 a 25034 c 24818 g 28459 t 1500 others
ORIGIN
Query Match 26.5%; Score 105; DB 2; Length 108765;
Best Local Similarity 66.7%; Pred. No. 3.8e-23;
Matches 150; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Qy 1 atcaagagtacagcctgaccagctgcccgcctatcgcaatgtgtggagccctctgg 60
Db 68859 ATGCAGAGTTTGAGTCCAGATCCCAAGCCCGATGACACAGTGTCTACGAGCCCTCAAG 68918
Qy 61 aggtattataagacgaggcctatgagccctatgagggcctgaaacgctcacagcaaca 120

```

Db 68919 AAAATCATCGGACGAGGCTCTGGAGCCCTTGGAGCGCTCAACGTCATGATCATG 68978
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy 121 ggcgcaggccctgccacgccctttatttgcctgctacgaaaagttaaaaaagacattg 180
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 68979 GGTGCAGGCGCCAGCCATGCTATTTTGGCTGCTATGAAACATGAAAAGGACTTTA 69038
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatggt 225
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 69039 AATGACGTTTTCACCAAGCAAGAAACAGCCACCTAGCCACGGT 69083
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 15
AX306258
LOCUS AX306258 483 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 1009 from Patent WO0188188.
ACCESSION AX306258
VERSION AX306258.1 GI:17645513
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 1009 22-NOV-2001;
          School Juridical Person Nihon University (JP)
FEATURES
          Location/Qualifiers
            source
              1..483
                /organism="Mus musculus"
                /db_xref="taxon:10090"
BASE COUNT 127 a 139 c 121 g 96 t
ORIGIN

Query Match 26.5%; Score 104.8; DB 6; Length 483;
Best Local Similarity 94.0%; Pred. No. 1.8e-23;
Matches 109; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 165 gtaaaaaagacattgagtgatgtaataccacctgggggcaatagccatattgccaatgg 224
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 15 GTTAAAGAGACATTGAGTGACGTAATCCACCCAGGGGGCAATAGCCATATGCCAATGG 74
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy 225 tgcggccgggtgtgtgcaacattacttcattgatgcagccatgaacccctgcggaag 280
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 75 TGCAGCCGGATGTGTGGGACATTACTTCATGATGCAGCCATGAATCCAGCGGAAG 130
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

Search completed: September 28, 2002, 08:00:04
Job time: 14925 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 08:06:28 ; Search time 664.31 Seconds
(without alignments)
1023.465 Million cell updates/sec

Title: US-09-870-113-11
Perfect score: 396
Sequence: 1 atgcagagtacagctga.....ttacctgtttaccactag 396

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	281	71.0	1418	21	Human ORF398
2	280	70.7	1294	22	Human polynucleoti
3	280	70.7	1322	22	Human transport pr
4	280	70.7	1336	21	Human uncoupling p
5	273	68.9	401	22	Novel human polynu
6	272	68.7	1316	22	Human polynucleoti
7	225	56.8	2502	22	Human secreted pro
8	225	56.8	27960	22	Human immune/haema
9	225	56.8	27960	22	Human immune/haema

10	174	43.9	339	22	Human immune/haema
11	129.6	32.7	1902	22	Human polynucleoti
12	128	32.3	1305	22	Human mitochondria
13	128	32.3	1716	21	Human ORF2744
14	128	32.3	2562	22	Human bone marrow
15	121	30.6	300	20	Human gene express
16	121	30.6	710	20	Human bone marrow
17	107.6	27.2	1913	22	Human ORF2730
18	105	26.5	842	22	Human cDNA 5'-end
19	105	26.5	842	22	Human cDNA clone r
20	105	26.5	1814	22	Human full-length
21	105	26.5	2837	22	Human immune/haema
22	104.8	26.5	483	24	Mouse ischaemic co
23	103.4	26.1	1890	22	Human full-length
24	103.4	26.1	3028	22	Human polynucleoti
25	101.8	25.7	1468	21	Human ORF2730
26	101	25.5	622	22	Human cDNA 5'-end
27	101	25.5	622	22	Human cDNA clone r
28	91.4	23.1	1758	20	Human secreted pro
29	91.4	23.1	1758	22	Human secreted pro
30	58	14.6	2037	24	cDNA sequence #574
31	42.2	10.7	2137	23	Drosophila melanog
32	37.4	9.4	8211	22	Human immune/haema
33	35	8.8	3842	22	cDNA encoding nove
34	34.4	8.7	6712	21	Human ORF2246
35	34	8.6	673	21	Aspergillus oryzae
36	33.6	8.5	1371	21	Arabidopsis thalia
37	33.6	8.5	1373	21	Arabidopsis thalia
38	33.6	8.5	11800	22	Human immune/haema
39	33.2	8.4	1227	23	DNA encoding novel
40	32.4	8.2	1311	18	Blackcurrant priB7
41	32.4	8.2	5150	18	Blackcurrant fruit
42	32.4	8.2	6326	23	Drosophila melanog
43	32	8.1	355	22	Streptococcus muta
44	32	8.1	1114	24	Human mitochondria
45	32	8.1	1363	20	cDNA encoding a pr

ALIGNMENTS

RESULT 1
AAC74843
ID AAC74843 standard; cDNA: 1418 BP.
XX
AC AAC74843;
XX
DT 08-FEB-2001 (first entry)
XX
XX Human ORF398 polynucleotide sequence SEQ ID NO:795.

Human: open reading frame; ORF3: detection; cytostatic; hepatotropic;
vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antiinflammatory;
antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive; ss.

Homo sapiens.

WO2000058473-A2.

05-OCT-2000.

31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX (CURA-) CURAGEN CORP.
 XX Shmkets RA, Leach M;
 PI
 XX
 PI
 XX
 PI
 XX
 DR WPI; 2000-602362/57.
 DR P-PSDB; AAB40634.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 5; Page 853-854; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiproliferative; antiparinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 1418 BP; 353 A; 400 C; 368 G; 297 T; 0 other;

Query Match 71.0%; Score 281; DB 21; Length 1418;
 Best Local Similarity 100.0%; Pred. No. 2.9e-85;
 Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 atgcagagctctacagctgaccagctgcccgcctatgcgaatgtgttgagccctctgg 60
 Db 178 atgcagagctctacagctgaccagctgcccgcctatgcgaatgtgttgagccctctgg 237
 Qy 61 aggtattataaagaacggggcctatgagggccatgaggggctgaacgtcacagcaaca 120
 Db 238 aggtattataaagaacggggcctatgagggccatgaggggctgaacgtcacagcaaca 297
 Qy 121 ggcgcagggcctgcccagcccttattttgcctgctacgaaaaagttaaaaaacacattg 180
 Db 298 ggcgcagggcctgcccagcccttattttgcctgctacgaaaaagttaaaaaacacattg 357
 Qy 181 agtgatgtaattccaccctgggggccaatagccatattgccaatggtgcggcggtgtgtg 240
 Db 358 agtgatgtaattccaccctgggggccaatagccatattgccaatggtgcggcggtgtgtg 417
 Qy 241 gcaacattacttactgatgacgcatgaacccctgcggaag 281
 Db 418 gcaacattacttactgatgacgcatgaacccctgcggaag 458

RESULT 2

AAI60661/C

ID AAI60661 standard; cDNA; 1294 BP.

XX

AC AAI60661;
 XX 22-OCT-2001 (first entry)
 DT
 XX Human polynucleotide SEQ ID NO 4650.
 DE
 XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 XX Homo sapiens.
 OS
 XX W0200153312-A1.
 PN
 XX 26-JUL-2001.
 PD
 XX 26-DEC-2000; 2000WO-US34263.
 PF
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang V, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM41505.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 XX Claim 1; SEQ ID NO 4650; 10078pp; English.
 PS
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nontropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 1294 BP; 267 A; 335 C; 368 G; 324 T; 0 other;
 Query Match 70.7%; Score 280; DB 22; Length 1294;
 Best Local Similarity 100.0%; Pred. No. 6.1e-85;
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 atgcagagctctacagctgaccagctgcccgcctatgcgaatgtgttgagccctctgg 60
 Db 1134 ATGCAGAGCTCTACAGCTGACCCAGCTGCCCGCTATCGCATGTGTGGAGGCCCTCTGG 1075
 Qy 61 aggtattataaagaacggggcctatgagggccatgaggggctgaacgtcacagcaaca 120


```
Db 1074 AGGATTATAGAACGAGGCGCTATGGAGCCCATGAGGGGCTGAACGTACAGCAACA 1015
QY 121 ggcgcaggcccgccacgccctttatttgcctgctacgaaagttaaaagacattg 180
Db 1014 GCGCAGGCGCTGCCACGCCCTTATTTTGGCTGCTAGCAAAAGTTAAAGACATTG 955
QY 181 agtcatgtaataccacccctggggccaatagcccatattgccaatggcgccgggtgtg 240
Db 954 AGTGATGTAATCCACCTGGGCAATAGCCATATTGCCAATGGTGCGCCGGGTGTG 895
QY 241 gcaacattactcatgatcagccatgaaccctgcggaag 280
Db 894 GCAACATTACTTCATGATGACCCATGAACCCCTGCGGGAAG 855

RESULT 3
AAF27733
ID AAF27733 standard; cDNA; 1322 BP.
XX AC AAF27733;
XX DT 28-MAR-2001 (first entry)
XX DE Human transport protein TPPT-33 coding sequence.
XX KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer; ss.
XX OS Homo sapiens.
XX PN WO200078953-A2.
XX PD 28-DEC-2000.
XX PF 16-JUN-2000; 2000WO-US16668.
XX PR 17-JUN-1999; 99US-0139923.
XX PR 10-AUG-1999; 99US-0148177.
XX PR 18-AUG-1999; 99US-0149357.
XX PR 28-OCT-1999; 99US-0162287.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX WPI: 2001-041424/05.
XX DR P-PSDB; AAB60113.
XX PT Isolated polypeptide with a human transport protein sequence is useful
XX PT for the diagnosis, prevention and treatment of disorders associated
XX PT with the immune, reproductive and cardiovascular systems -
XX PS Claim 5; Page 160; 165pp; English.
XX CC The present invention provides the protein and coding sequences for 43
XX CC novel human transport proteins (designated TPPTs). These can be used in
XX CC the diagnosis and treatment of transport, metabolic, neurological,
XX CC reproductive, cardiovascular and immune disorders, and cell proliferative
XX CC disorders such as cancer.
XX SQ Sequence 1322 BP; 328 A; 376 C; 351 G; 267 T; 0 other;

Query Match 70.7%; Score 280; DB 22; Length 1322;
Best Local Similarity 100.0%; Pred. No. 6.2e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagtctacagctgacccagctgcccgctatcgcaatgtgttgaggccctctgg 60
Db 186 atgcagagtctacagcctgacccagctgcccgctatcgcaatgtgttgaggccctctgg 245
```

```
QY 61 aggattataaagcagggagccctatgagggcccatgagggcctgaacgtcacagcaaca 120
Db 246 aggattataaagcagggagccctatgagggcccatgagggcctgaacgtcacagcaaca 305
QY 121 ggcgcaggccctgcccacgccctttatttgcctgctacgaaagttaaaagacattg 180
Db 306 ggcgcaggccctgcccacgccctttatttgcctgctacgaaagttaaaagacattg 365
QY 181 agtcatgtaataccacccctggggccaatagcccatattgccaatggcgccgggtgtg 240
Db 366 agtcatgtaataccacccctggggccaatagcccatattgccaatggcgccgggtgtg 425
QY 241 gcaacattactcatgatcagccatgaaccctgcggaag 280
Db 426 gcaacattactcatgatcagccatgaaccctgcggaag 465

RESULT 4
AAC90457
ID AAC90457 standard; cDNA; 1336 BP.
XX AC AAC90457;
XX DT 12-MAR-2001 (first entry)
XX DE Human uncoupling protein cDNA #6.
XX KW Human; uncoupling protein; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cardiant; vasotropic;
KW cerebroprotective; neuroprotective, antibacterial; ophthalmological;
KW gastrointestinal; nephrotropic; gynaecological; vulnery; thrombolytic;
KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
KW infertility; ss.
XX OS Homo sapiens.
XX PN WO200061614-A2.
XX PD 19-OCT-2000.
XX PF 06-APR-2000; 2000WO-US09534.
XX PR 09-APR-1999; 99US-0128701.
XX PR 08-JUL-1999; 99US-0142821.
XX PR 18-AUG-1999; 99US-0149448.
XX PR 12-NOV-1999; 99US-0164751.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;
XX WPI: 2000-656322/63.
XX DR P-PSDB; AAB50383.
XX PT Uncoupling proteins and nucleic acid sequences encoding them, useful
XX PT for detecting, preventing and treating proliferative, neurological,
XX PT immune system, cardiovascular and gastrointestinal disorders -
XX PS Claim 1; Page 306-307; 343pp; English.
XX CC The present sequence is one of eighteen isolated nucleotide sequences
XX CC encoding uncoupling proteins. The nucleotide sequences may be used for
XX CC the detection of various disorders such as cancer, for chromosome
XX CC identification, as chromosome markers and for numerous other diagnostic
XX CC or research purposes. The uncoupling protein encoded by the nucleotide
XX CC sequences may be used to treat disorders such as neural, immune,
XX CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
XX CC renal and proliferative disorders, wounds, infectious diseases,
XX CC thrombosis, arthritis, and infertility.
XX SQ Sequence 1336 BP; 333 A; 377 C; 358 G; 268 T; 0 other;
```

```
Query Match      70.7%; Score 280; DB 21; Length 1336;
Best Local Similarity 100.0%; Pred. No. 6.2e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgacccagctcccgctatcgaatgtgttgagccctctgg 60
    |||
Db 197 atgcagagctacagcctgacccagctcccgctatcgaatgtgttgagccctctgg 256

Qy 61 aggattataagaacgagggcctatggagcccatgagggcctgaacgtcacagcaaca 120
    |||
Db 257 aggattataagaacgagggcctatggagcccatgagggcctgaacgtcacagcaaca 316

Qy 121 ggcgcagggcctgccacgcgccttattttgctgctacgaaagttaaaaaagacattg 180
    |||
Db 317 ggcgcagggcctgccacgcgccttattttgctgctacgaaagttaaaaaagacattg 376

Qy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatgtgcggcggtgtgtg 240
    |||
Db 377 agtgaatgaatccaccctgggggcaatagccatattgccaatgtgcggcggtgtgtg 436

Qy 241 gcaacattacttcattgatgcagccatgaacctgcggaag 280
    |||
Db 437 gcaacattacttcattgatgcagccatgaacctgcggaag 476

RESULT 5
AAF66156
ID AAF66156 standard; cDNA; 401 BP.
XX
AC AAF66156;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1912.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klingler J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI KrakenJakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
PS Claim 9; Page 818; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
```

```
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match      68.9%; Score 273; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 8.2e-83;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 gtctacagcctgacccagctcccgctatcgaatgtgttgagccctctggaggatta 67
    |||
Db 9 gtctacagcctgacccagctcccgctatcgaatgtgttgagccctctggaggatta 68

Qy 68 taagaacgagggcctatggagggcccatgaggggctgaacgtcacagcagcgag 127
    |||
Db 69 taagaacgagggcctatggagggcccatgaggggctgaacgtcacagcagcgag 128

Qy 128 ggcctgccacgcgccttattttgctgctacgaaaagttaaaaaagacattgagtgtg 187
    |||
Db 129 ggcctgccacgcgccttattttgctgctacgaaaagttaaaaaagacattgagtgtg 188

Qy 188 taatccaccctgggggcaatagccatattgccaatgtgcggcggtgtgtggcaacat 247
    |||
Db 189 taatccaccctgggggcaatagccatattgccaatgtgcggcggtgtgtggcaacat 248

Qy 248 tacttcattgatgcagccatgaacctgcggaag 280
    |||
Db 249 tacttcattgatgcagccatgaacctgcggaag 281

RESULT 6
AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX
AC AAI58875;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1078.
XX
KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
```

```
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB: AAM39719.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 1078; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;

Query Match 68.7%; Score 272; DB 22; Length 1316;
Best Local Similarity 98.2%; Pred. No. 3.3e-82;
Matches 275; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgacccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
Db 176 atgcagtctacagctgacccagctgcccgcctatgcgaatgtgttgaggctcctctgg 235
Qy 61 aggtattataagaacgaggccctatgagcccatgagggcctggaacgtcacagcaaca 120
Db 236 aggtattataagaacgaggccctatgagcccatgagggccatgagggccgcaacagcaaca 295
Qy 121 ggcgcagggcctgccagcccttatttgcctgctacgaaaaagttaaaaaagacattg 180
Db 296 ggcgcagggcctgccagcccttatttgcctgctacgaaaaagttaaaaaagacattg 355
Qy 181 agtgatgtaattccacctgggggcaataagccatattgccatgtgtgcggcggtgtgtg 240
Db 356 agtgatgtaattccacctgggggcaataagccatattgccatgtgtgcggcggtgtgtg 415
Qy 241 gcaacattacttcatgctgcagcccatgaacctgcggaag 280
Db 416 gcaacattacttcatgctgcagcccatgaacctgcggaag 455

RESULT 7
AAS03906
ID AAS03906 standard; cDNA; 2502 BP.
XX
XX AAS03906;
XX
XX 29-AUG-2001 (first entry)
XX
XX Human secreted protein gene #25.
XX
XX Human secreted protein; autoimmune disorder; hyperproliferative disorder;
XX cardiovascular disorder; cerebrovascular disorder; angiogenesis;
XX nervous system disorder; bacterial infection; viral infection; ss;
XX fungal infection; ocular disorder; wound healing; tissue regeneration;
XX epithelial cell proliferation; skin ageing; chemotaxis; I9G Fc region.
XX

OS Homo sapiens.
XX
XX WO200123598-A1.
XX
XX 05-APR-2001.
XX
XX 26-SEP-2000; 2000WO-US26324.
XX
XX 27-SEP-1999; 99US-0155807.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Komatsoulis G, Ruben SM, Rosen CA;
XX
XX WPI: 2001-281684/29.
XX
XX P-PSDB: RAU01950, RAU01989.
XX
XX Forty one nucleic acid molecules encoding human secreted proteins, useful
XX in the prevention, treatment and diagnosis of cancer, immune disorders,
XX cardiovascular disorders and neurological diseases -
XX
XX Disclosure; Page 460-461; 518pp; English.
XX
XX Sequences AAS03873-AAS03922 represent isolated nucleic acid molecules and
XX PCR primers of the invention. acid of the invention. Secreted proteins
XX and their related nucleic acids can be used in the diagnosis of or
XX susceptibility to a pathological condition by determining the presence or
XX absence of a mutation in a nucleic acid or the presence or amount of
XX expression of a secreted protein. The sequences are used to prevent,
XX treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
XX goats, horses, cats, dogs, chickens or sheep. The antibodies to the
XX polypeptides can also be used in alleviating symptoms associated with
XX disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
XX linked immunosorbent assays (ELISA). The disorders include autoimmune
XX diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
XX neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
XX arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
XX nervous system disorders e.g. Alzheimer's disease, infections caused by
XX bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
XX The peptides can also be used to aid wound healing and epithelial cell
XX proliferation, to help prevent skin ageing due to sunburn, to maintain
XX organs before transplantation, to regenerate tissues, in chemotaxis and
XX as a food additive or preservative to alter storage capabilities.
XX
XX Sequence 2502 BP; 587 A; 655 C; 669 G; 589 T; 2 other;

Query Match 56.8%; Score 225; DB 22; Length 2502;
Best Local Similarity 100.0%; Pred. No. 5e-66;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgacccagctgccgcctatgcgaatgtgttgaggccctctgg 60
Db 195 atgcagagtctacagcctgacccagctgccgcctatgcgaatgtgttgaggccctctgg 254
Qy 61 aggtattataagaacgaggccctatgagcccatgagggcctggaacgtcacagcaaca 120
Db 255 aggtattataagaacgaggccctatgagcccatgagggcctggaacgtcacagcaaca 314
Qy 121 ggcgcagggcctgccagcccttatttgcctgctacgaaaaagttaaaaaagacattg 180
Db 315 ggcgcagggcctgccagcccttatttgcctgctacgaaaaagttaaaaaagacattg 374
Qy 181 agtgatgtaattccacctgggggcaataagccatattgccatgtgttccaatggt 225
Db 375 agtgatgtaattccacctgggggcaataagccatattgccatgtgttccaatggt 419

RESULT 8
AAK69779
ID AAK69779 standard; DNA; 27960 BP.
XX
XX AAK69779;
XX
```

XX 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24591.
XX
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
KW
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205513.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-02331968.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.
XX 14-SEP-2000; 2000US-0232399.
XX 14-SEP-2000; 2000US-0232400.
XX 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis.
XX
XX
PS Disclosure; SEQ ID NO 24591; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;

Query Match 56.8%; Score 225; DB 22; Length 27960;
Best Local Similarity 100.0%; Pred. No. 1.7e-65;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
Db 18295 atgcagagtctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 18354
|||||

Qy 61 aggattataagaacagaggagccctatggagcccatgaggggctgaacgtcacagcaaca 120
|||||
Db 18355 aggattataagaacagaggagccctatggagcccatgaggggctgaacgtcacagcaaca 18414
|||||

Qy 121 ggcgcaggcctgccacgcctctatttgcctctacgaaaagttaaaaaagacattg 180
|||||
Db 18415 ggcgcaggcctgccacgcctctatttgcctctacgaaaagttaaaaaagacattg 18474
|||||

Qy 181 agtgatgtaataccacctgggggcaatgccaatgttgcgaatggt 225
|||||
Db 18475 agtgatgtaataccacctgggggcaatgccaatgttgcgaatggt 18519
|||||

RESULT 9
AAK73320
ID AAK73320 standard; DNA; 27960 BP.
XX
XX
AC AAK73320;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28132.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
KW

XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354..
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.

PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0233423.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX P-PSDB; AAM85891.

PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -
PT	
PT	
XX	Claim 1; SEQ ID NO 3732; 3071pp + Sequence Listing; English.
XX	
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK34942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention.
XX	
SO	Sequence 339 BP; 86 A; 84 C; 101 G; 64 T; 4 other;
	Query Match 43.9%; Score 174; DB 22; Length 339;
	Best Local Similarity 94.9%; Pred. No. 4.9e-49;
	Matches 188; Conservative 1; Mismatches 8; Indels 1; Gaps 1
QY	1 atgcagagctctacagcctgaccagctcccgctatcgcaatgtgttgaggccctctgg 60 : : : : : : : : :
Db	143 atcagwgctctacagcctgaccagctcccgctatcgcaatgtgttgaggccctctgg 202
QY	61 aggattataagaacggaggggccctatggaggcccatgagggggctgaacgtcacagcaaca 120 : : : : : : : : :
Db	203 aggattataagaacggaggggccctatggagg-ccatgagggggctgaacgtcacagcaaca 261
QY	121 ggcgcaggccctccccacgcctttatttgcctgctacgaaaagttaaaaagacattg 180 : : : : : : : : :
Db	262 ggcgcaggggcccnccacgcgctttatttgcctgtacggaaaagttcaaaaaggcattg 321
QY	181 agtgatgtaatccaccct 198 : : : : : : : : :
Db	322 agtgatgtaatccaccct 339
RESULT	11
AAK52172	
ID	AAK52172 standard; cDNA; 1902 BP.
XX	
AC	AAK52172;
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human polynucleotide SEQ ID NO 717.
XX	
XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; hematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation; ss.
OS	Homo sapiens.
XX	
PN	WO200157190-A2.
XX	
PD	09-AUG-2001.
XX	
XX	05-FEB-2001; 2001WO-US04098.
PF	
XX	
PR	03-FEB-2000; 2000US-0496914.
PR	27-APR-2000; 2000US-0560875.

XX Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
 KW preparation; detection; ss.
 XX Homo sapiens.
 OS CN1269409-A.
 XX 11-OCT-2000.
 XX 17-MAR-2000; 2000CN-0114958.
 XX 17-MAR-2000; 2000CN-0114958.
 XX (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
 PA Zhang X, Gao X, Xiao H;
 XX WPI; 2001-050544/07.
 DR P-PSDB; AAB60658.
 XX New human mitochondrion solute carrier protein and its nucleic acid -
 XX Claim 1; Page 19-20; 21pp; Chinese.
 CC The invention relates to a novel human mitochondrial solute carrier
 CC protein, hMSC-o (AAB60658), and cDNA encoding it (AAF59920). hMSC-o is
 CC expressed in normal human hypothalamus tissue. The invention also relates
 CC to the preparation of hMSC-o proteins and nucleic acids, and the
 CC detection of hMSC-o proteins and nucleic acids in a sample. The present
 CC sequence represents cDNA encoding hMSC-o.
 XX Sequence 1305 BP; 292 A; 388 C; 363 G; 262 T; 0 other;
 SQ

Query Match 32.3%; Score 128; DB 22; Length 1305;
 Best Local Similarity 66.1%; Pred. No. 4.5e-33;
 Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1 atgcagagttcagcagctgaccagctgccgctatcgcaatgtgtgagccctctg 60
 Db 172 atgcagagtttgatcagatcccaagccagtcacacagtcacagccctcaag 231
 QY 61 aggattataagacaggggctctgagggccatgaggggctgaacgtcacagcaaca 120
 Db 232 aaatcatcgagcagagctctgagggccttgagggcgtcaacgtcatgctatg 291
 QY 121 ggcgagggcctgccagccctttatttctgctgtacgaaagttaaaagacattg 180
 Db 292 ggtcgaggccgcccattgctgtatttctgctgtatgctgtatgaaacatgaaaggactta 351
 QY 181 agtgatgtaatccaccctggggcctatgagccatgattgccaatggtgcggcggtgtgtg 240
 Db 352 atagcgttttcaccacacacagaaacacccacccacccacacagggatagctgggagatg 411
 QY 241 gcaacattactctatgatcagccatgaacctggggaag 280
 Db 412 gccacctgtccagatcggtgtaataatgaatccagcagaag 451

RESULT 13
 AAC77189
 ID AAC77189 standard; cDNA; 1716 BP.
 AC AAC77189;
 XX 08-FEB-2001 (first entry)
 XX Human ORFX ORF2744 polynucleotide sequence SEQ ID NO:5487.
 DE Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
 KW vulnary; antiprosaric; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;

KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX Homo sapiens.
 XX WO200058473-A2.
 XX 05-OCT-2000.
 XX 31-MAR-2000; 2000WO-US08621.
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX (CURA-) CURAGEN CORP.
 PA Shinkets RA, Leach M;
 PI WPI; 2000-602362/57.
 DR P-PSDB; AAB42980.
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX Claim 5; Page 4661-4662; 5507pp; English.
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiprosaric; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antirheumatic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX Sequence 1716 BP; 411 A; 457 C; 446 G; 402 T; 0 other;
 SQ

Query Match 32.3%; Score 128; DB 21; Length 1716;
 Best Local Similarity 66.1%; Pred. No. 5.2e-33;
 Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1 atgcagagttcagcagctgaccagctgccgctatcgcaatgtgtgagccctctg 60
 Db 256 atgcagagtttgatcagatcccaagccagtcacacagtcacagccctcaag 315
 QY 61 aggattataagacaggggctctgagggccatgaggggctgaacgtcacagcaaca 120

Db 316 aaatcatcgagaccgaagctttctgtgagcccttgagcgtcgaacgtcatgcatg 375
Qy 121 ggcagagcctccacgcctttatttgcctgtcacgaaagttaaaagacattg 180
Db 376 ggtgagggcagcccatgcatgtatttgcctgtcgaacatgaaagacttta 435
Qy 181 agtgatgaatccaccctggggcgaatagccatatttcccaatggtgcgcgggtgtg 240
Db 436 aatgacgttttccaccacgaagaaacagccaccctagccaacggatagctgggagtg 495
Qy 241 gcaacattactatgatgacgacatgaacccctgcggaag 280
Db 496 gccacctgtcccagatgctgtaagtgaatccacagagaag 535

RESULT 14
AAH90057
ID AAH90057 standard; cDNA; 2562 BP.
XX
AC AAH90057;
XX
DT 01-OCT-2001 (first entry)
XX Human bone marrow cDNA, SEQ ID NO: 301.
DE
XX
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
OS Homo sapiens.
XX
XX WO200153453-A2.
XX
XX 26-JUL-2001.
XX
XX 23-DEC-2000; 2000WO-US34960.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX
XX WPI; 2001-488707/53.
DR P-PSDB; AAM00938.
XX
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1; Page 412-413; 648pp; English.
XX
XX The present sequence is one of 251 novel human polynucleotides
CC expressed in the bone marrow. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one

CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX
SQ Sequence 2562 BP; 623 A; 660 C; 627 G; 652 T; 0 other;

Query Match 32.3%; Score 128; DB 22; Length 2562;
Best Local Similarity 66.1%; Pred. No. 6.4e-33;
Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagctgacccagctgcgcgtatcgcgaatgtgttgagccctctg 60
Db 2021 atgcagagtttgatccagatcccaagcccatgacacagttactacgagccctcaag 2080
Qy 61 aggtattataagaacgagggccttatgagcccatgaggggctgaacgtcacgcaaca 120
Db 2081 aaatcatcgagaccgaagcttctgagcccttgagcgtcgaacgtcatgcatg 2140
Qy 121 ggcagagggcctgccacgcctttatttgcctgtcacgaaagttaaaagacattg 180
Db 2141 ggtgagggcgcgcctatgcatgtatttgcctgtcgaacatgaaagacttta 2200
Qy 181 agtgatgaatccaccctggggcgaatagccatatttcccaatggtgcgcgggtgtg 240
Db 2201 aatgacgttttccaccacgaagaaacagccacccctagccaacggatagctgggagtg 2260
Qy 241 gcaacattactatgatgacgacatgaacccctgcggaag 280
Db 2261 gccacctgtcccagatgctgtaagtgaatccacagagaag 2300

RESULT 15
AAZ13026
ID AAZ13026 standard; cDNA; 300 BP.
XX
AC AAZ13026;
XX
DT 12-OCT-1999 (first entry)
XX Human gene expression product cDNA sequence SEQ ID NO:495.
DE
XX Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
XX WO9938972-A2.
XX
XX 05-AUG-1999.
XX
XX 28-JAN-1999; 99WO-US01619.
XX
XX 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX
XX WPI; 1999-494092/41.
DR
XX Novel human genes and their expression products which are
PT differentially expressed in different cell types

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:56:00 ; Search time 182.06 Seconds
(without alignments)
534.279 Million cell updates/sec

Title: US-09-870-113-11
Perfect score: 396
Sequence: 1 atgcagagctctacagcctga.....ttacctgtgtttaccactag 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32.4	8.2	1311	4	US-09-068-140A-9
2	32.4	8.2	5150	4	US-09-068-140A-14
3	32	8.1	1643	2	US-08-933-750C-68
4	32	8.1	1643	3	US-08-234-613-68
5	30.2	7.6	2277	1	US-08-676-967-2
6	30.2	7.6	2277	2	US-08-676-974-2
7	30.2	7.6	2277	2	US-09-098-487-2
8	29.6	7.5	4091	4	US-08-630-915A-33
9	29.2	7.4	5532	2	US-08-475-035-3
10	29	7.3	1747	1	US-08-202-389-1
c 11	29	7.3	8600	4	US-09-457-708-1
c 12	28.6	7.2	1311	3	US-09-073-569-4
c 13	28.4	7.2	4368	1	US-08-026-138E-17
c 14	28.4	7.2	4446	1	US-08-026-138E-6
c 15	28.2	7.1	7011	4	US-09-268-163-9
16	27.8	7.0	1323	4	US-09-297-937C-10
17	27.8	7.0	4665	4	US-09-297-937C-12
c 18	27.8	7.0	5643	1	US-08-144-602B-4
19	27.6	7.0	3942	4	US-09-162-484-19
c 20	27.2	6.9	3771	1	US-08-185-432-3
21	27.2	6.9	3771	1	US-08-185-432-23
22	27.2	6.9	3810	2	US-08-475-844-8
23	27.2	6.9	3810	5	PCT-US95-08429-8
c 24	27.2	6.9	5063	1	US-08-185-432-1
c 25	27	6.8	591	1	US-08-686-878A-5
c 26	27	6.8	2379	3	US-08-537-361E-5
c 27	27	6.8	3153	4	US-09-175-928-9

28	27	6.8	3508	4	US-08-462-457B-7	Sequence 7, Appli
29	27	6.8	3508	4	US-08-334-179A-7	Sequence 7, Appli
30	27	6.8	7653	4	US-08-471-112A-1	Sequence 11, Appli
31	27	6.8	7824	5	PCT-US95-06722-11	Sequence 15, Appli
c 32	26.8	6.8	1128	3	US-09-106-217-15	Sequence 1, Appli
c 33	26.8	6.8	1134	3	US-09-106-217-1	Sequence 1, Appli
34	26.8	6.8	1192	4	US-09-142-565-1	Sequence 38, Appli
35	26.8	6.8	1255	1	US-08-518-878B-38	Sequence 38, Appli
36	26.8	6.8	1255	1	US-08-294-522B-38	Sequence 38, Appli
37	26.8	6.8	1255	2	US-08-470-868A-38	Sequence 38, Appli
38	26.8	6.8	1596	2	US-08-807-861A-38	Sequence 38, Appli
39	26.8	6.8	1596	3	US-09-210-681-38	Sequence 38, Appli
40	26.8	6.8	1596	3	US-08-946-719A-38	Sequence 38, Appli
c 41	26.8	6.8	3672	4	US-09-056-105-17	Sequence 20, Appli
42	26.8	6.8	246240	2	US-08-724-394A-20	Sequence 21, Appli
43	26.8	6.8	246240	2	US-08-724-394A-21	Sequence 22, Appli
44	26.8	6.8	246240	2	US-08-724-394A-22	Sequence 138, App
45	26.6	6.7	573	2	US-08-290-665A-138	

ALIGNMENTS

RESULT 1
US-09-068-140A-9
; Sequence 9, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-9

Query Match	8.1%	Score 32;	DB 3;	Length 1643;
Best Local Similarity	49.4%;	Pred. No.	0.29;	
Matches	83;	Conservative	0;	Mismatches 85; Indels 0; Gaps 0;
Qy	26	ctgcgccgctatcgcaatgtgttgaggccctctggaggtattataagaacgagggccttat	85	
Db	482	CTACCCGGTTCACATGGGACCATTGCCTTCGTGAAGATCGTGAGGCCACGAGGGCACCA	541	

[illegible]

```

6
US-08-676-974-2
; Sequence 2, Application US/08676974
; Patent No. 5770422
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,974
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-676-974-2

Query Match 7.6%; Score 30.2; DB 1; Length 2277;
Best Local Similarity 27.4%; Pred. No. 1.5;
Matches 80; Conservative 50; Mismatches 162; Indels 0; Gaps 0;

Qy 86 ggagggccatgaggggctgaacgtcacagcaacagcgagcgccctgcccacgcccttt 145
Db 503 GNAARGCNYTNAARGGNATGAAYATGAARGARATHAARGGNMGACNGTNGCNGTNGAYT 562
Qy 146 atttgctgtacgaaagttaaaagacattgagtgatgtaatccaccctggggca 205
Db 563 GGCNGTNGCNAARGAAYATAYAAAGAYACNARWSNGTNGCNGTNGAYT 622
Qy 206 atagccatattgccaatggtgcggccggtgtgtggaacacattactcatgatgcagcca 265
Db 623 ARWSNCAYGARWSNAARCAAYCARGARWSNGTNAARAARAARGGNMGARGARGAYA 682
Qy 266 tgaacctcggaaggtaagtattcctcaacctatcactctgtggcagctgcacctgta 325
Db 683 TGGARGARGARGAARAAAYGAYGAYGAYGAYGAYGAYGAYGARGARGAYGGNGTNTTYG 742
Qy 326 ttctttacattgttcagaagaagcacatcagtttgggtgggaaattctgt 377
Db 743 AYGAYGARGAYGARGARGARGAARAYATHGARWSNAARGTNACNAARCCNGT 794

RESULT 6
US-08-630-915A-33
; Sequence 33, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
7
US-09-098-487-2
; Sequence 2, Application US/09098487
; Patent No. 5917025
; GENERAL INFORMATION:
```


;; TITLE OF INVENTION: USING SAME
;; NUMBER OF SEQUENCES: 227
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds LLP
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10036-2711
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/630,915A
;; FILING DATE: 03-APR-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MISTOCK, S. LESLIE
;; REGISTRATION NUMBER: 18,872
;; REFERENCE/DOCKET NUMBER: 1101-174
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 33:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4091 bases
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
US-08-630-915A-33

Query Match 7.5%; Score 29.6; DB 4; Length 4091;
Best Local Similarity 51.5%; Pred. No. 3.2;
Matches 68; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 10 ctacagctgacccagctgcccgtatgcaatggtgtgagggccctctggaggattata 69
Db 995 CGAGACTGGAGCAGGCCATCCGGGGCCGATGCCAGGAGGACCTCAGGTGGTCCGC 1054
Qy 70 agaacggaggccctatggaggcccatgaggggctgaactcacagcaacagcgccaggg 129
Db 1055 AGCACAGTGGCCCGGATGCCATGAACCTGCGCCAGTTCGAGAGTGGAAACCAGAC 1114
Qy 130 cctgcccacgccc 141
Db 1115 CTCCCGCACACC 1126

RESULT 9
US-08-475-035-3
; Sequence 3, Application US/08475035
; Patent No. 5985553
; GENERAL INFORMATION:
; APPLICANT: KING, C. R.
; APPLICANT: KRAUS, MATTHIAS H.
; APPLICANT: AARONSON, STUART A.
; TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
; TITLE OF INVENTION: EGF RECEPTOR GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/475,035
;; FILING DATE: 7 Jun 1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: PERRYMAN, DAVID G.
;; REGISTRATION NUMBER: 33,438
;; REFERENCE/DOCKET NUMBER: 1414.656
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 404/688-0770
;; TELEFAX: 404/688-9880
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5532 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 187..3816
US-08-475-035-3

Query Match 7.4%; Score 29.2; DB 2; Length 5532;
Best Local Similarity 53.5%; Pred. No. 5.2;
Matches 61; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 258 tgagccatgaaccctgagggaagtaattccctcaacctatcacctctgtggcgagctg 317
Db 4605 TGCTGTGATGAATCAGCAAGAGAGAGATGACACATCAATAATACTCGGATTCAGCCC 4664
Qy 318 cacctgtatttcttacagtttgcagaagaaagcacatcacgttttgggtggaaa 371
Db 4665 ACATTGGATTTCAGCATTTGGACCAATAGCCACAGCTGAGAATGTGGAATA 4718

RESULT 10
US-08-202-389-1
; Sequence 1, Application US/08202389
; Patent No. 5536636
; GENERAL INFORMATION:
; APPLICANT: FREEMAN JR., ROBERT M.
; APPLICANT: PLUTZKY, JORGE
; APPLICANT: NEEL, BENJAMIN G.
; APPLICANT: ROSENBERG, ROBERT D.
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
; TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,389
; FILING DATE: 28-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,926
; FILING DATE: 01-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,141

;; FILING DATE: 31-JAN-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/721,112
;; FILING DATE: 26-JUN-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahop, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: BH92-05WA
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 861-6240
;; TELEFAX: (617) 861-9540
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1747 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 2..1540
US-08-202-389-1

Query Match 7.3%; Score 29; DB 1; Length 1747;
Best Local Similarity 57.0%; Pred. No. 3.3;
Matches 53; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
Qy 96 gagggggtgaacgtcacagcaaacaggcgagggcctgccacgcccttttttgcgtg 155
Db 105 GAGGCGAGCGAGGTCATGCTGCGAGGCAAGGGCGAGCCCTGGACATTTCTGTGCGTG 164
Qy 156 ctacgaaagttaaaaaagacattgagtgatg 188
Db 165 AGAGTCTCAGCAACCTGCTGATTTTGTCCT 197

RESULT 11
US-09-457-708-1/c
; Sequence 1, Application US/09457708
; Patent No. 6326483
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, David J.
; APPLICANT: Sampson, Julian R.
; APPLICANT: Povey, Sue
; APPLICANT: van Slegtenhorst, Marjon
; APPLICANT: Halley, Dicky
; TITLE OF INVENTION: Compositions and Methods Based U
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Vinson & Elkins
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20004-1008
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/457,708
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanzo, Michael A.
; REGISTRATION NUMBER: 36,912
; REFERENCE/DOCKET NUMBER: BRI331/42002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 639-6585
; TELEFAX: (202) 639-6604

;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8600 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-09-457-708-1

Query Match 7.3%; Score 29; DB 4; Length 8600;
Best Local Similarity 58.8%; Pred. No. 7.7;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
Qy 100 gggctgaacgtcacagcaaacaggcgagggcctgccacgcccttttttgcctgctac 159
Db 7750 GGGACAAGGGTCACAGCAGCAGCCTAGGGGCGCAGTCTCGACCTAAACTTGTCTTTCAC 7691
Qy 160 gaaaagttaaaaaagacattgagtg 184
Db 7690 CTACAGACAAAAGCTTAATCAAGTG 7666

RESULT 12
US-09-073-569-4/c
; Sequence 4, Application US/09073569
; Patent No. 6084088
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: NOVEL TUMOR ANTIGENS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,569
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 97-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-073-569-4

Query Match 7.2%; Score 28.6; DB 3; Length 1311;
Best Local Similarity 30.7%; Pred. No. 3.9;
Matches 42; Conservative 25; Mismatches 70; Indels 0; Gaps 0;

Qy	170	aaagacattgagatgataatccaccctgggggcaatgaccatattgccaatgttgagg	229
		: : : : : : : : : : : : : : : : : :	
Db	626	ARRAARTCYTTNCKRTTRCARTTTCNGTCATCCNACNNGGCCDATYTCYTGNGTNCRC	567
Qy	230	ccgggtgtgtggcaacattactctatgatcgagccatgaacctgcgggaagtgaatgatt	289
		: : : : : : : : : : : : : : : : : : : :	
Db	566	TTNARNARRTRCANCNGGYTGNGGCATRCANCCTGNACNCKNARRTTSNRAADATN	507
Qy	290	cctcaacacctatcaactct	306
		: : : : : : : : : : : : : : : : :	
Db	506	CCNCCNCCNCKNARNC	490

```

RESULT 13
US-08-026-138E-17/c
; Sequence 17, Application US/08026138E
; Patent No. 5502166
; GENERAL INFORMATION:
; APPLICANT: Masayoshi MISHINA
; TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME

```

NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nishionata Residence 1-107
STREET: 5214, Nishionata-machi
CITY: Niigata-shi
STATE: Niigata-ken
COUNTRY: JAPAN
ZIP: 951
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS v.5
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/026.138E
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 39563/1992
FILING DATE: 26-FEB-1992
APPLICATION NUMBER: JP 173155/1992
FILING DATE: 30-JUN-1992
APPLICATION NUMBER: JP 215017/1992
FILING DATE: 12-AUG-1992
APPLICATION NUMBER: JP 303878/1992
FILING DATE: 13-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hamburg, C.Bruce
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-4551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 986-2340
TELEFAX: (212) 953-7733
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 4368 nucleic acids
TYPE: nucleic acid
STRANDEDNESS: double strand
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: mouse
TISSUE TYPE: brain
PUBLICATION INFORMATION:
AUTHORS: Masayoshi MISHINA
TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
RELEVANT RESIDUES IN SEQ ID NO: 17: FROM 1 TO 4368
PS-08-026-138E-17

```
Query Match      7.2%; Score 28.4; DB 1; Length 4368;
Best Local Similarity 70.4%; Pred. No. 8.7;
Matches 38; Conservative 0; Mismatches 16; Indels 0
```

Qy 64 attataaacgagggccctatgaggagcccatgaggggctgaacgtcacgca 117
||||| | ||| | |||| | |||| | |||| | |||| | |||| | ||||
Db 323 ATTATCATAGATGAGCCCGCCATGGATGCCAGGATGGGGTGTGAGAGTCTGAGCA 270

RESULT 14
US-08-026-138E-6/c
; Sequence 6, Application US/08026138E
; Patent No. 5502166
; GENERAL INFORMATION:
; APPLICANT: Masayoshi MISHINA
; TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nishiohata Residence 1-107
; STREET: 5214, Nishiohata-machi
; CITY: Niigata-shi
; STATE: Niigata-ken
; COUNTRY: JAPAN
; ZIP: 951
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS v.5
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/026,138E
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 39563/1992
; FILING DATE: 26-FEB-1992
; APPLICATION NUMBER: JP 173155/1992
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: JP 215017/1992
; FILING DATE: 12-AUG-1992
; APPLICATION NUMBER: JP 303878/1992
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamburg, C.Bruce
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-4551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 986-2340
; TELEFAX: (212) 953-7733
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4446 nucleic acids
; TYPE: nucleic acid
; STRANDEDNESS: double strand
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: mouse
; TISSUE TYPE: brain
; PUBLICATION INFORMATION:
; AUTHORS: Masayoshi MISHINA
; TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 4446
US-08-026-138E-6

```
Query Match          7.2%; Score 28.4; DB 1; Length 4446;
Best Local Similarity 70.4%; Pred. No. 8.8;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

Qy 64 attatagaacgaggggcctatggagggcccatgatgggggtgaactcacagca 117
||||| | ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 401 ATTATCATAGATGAGCCCCCATTGGATGCCAGGATGGGGGTGAGTCTGAGCA 348

RESULT 15
US-09-268-163-9/c

```
; Sequence 9, Application US/09268163B
; Patent No. 6353091
; GENERAL INFORMATION:
; APPLICANT: Lipscombe, Diane
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/09/268,163B
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 60/077,901
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 7011
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7008
US-09-268-163-9
```

```
Query Match      7.1%; Score 28.2; DB 4; Length 7011;
Best Local Similarity 54.3%; Pred. No. 13;
Matches 57; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 167 taaaaagacattgagtgtgtaattccaccctgggggcaatagccatattgccaatggtg 226
      ||||| ||| ||| ||| ||||| ||| ||||| ||| ||||| |||
Db 6357 TGAAGAGAGAGAGAGGGGCTCCCTCCGCTCTCTGGGACCGGCTCGCTCTGCTTACGCTC 6298

QY 227 cggccgggtgtgtggcaacattacttcattgatgcagccatgaacc 271
      ||||| ||| ||| ||| ||||| ||||| ||||| |||||
Db 6297 CCGCCGCGAGCCTGTGGACCCCTTCTCCATGGGGGCGAGCCAGATCC 6253
```

Search completed: September 28, 2002, 07:56:08
Job time: 11253 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 06:22:49 ; Search time 4881.77 Seconds
(without alignments)
1094.848 Million cell updates/sec

Title: us-09-870-113-11

Perfect score: 396

Sequence: 1 atcagagttacagctga.....ttacctgttttaccactag 396

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	280	70.7	747	10	BI827322 603077989
2	277.6	70.1	888	9	AL530804
3	265.8	67.1	480	10	BI045863 MR3-FN020
4	256	64.6	573	10	BI541703 455541 MA
5	246.4	62.2	598	10	BE913718 601669072
6	246.4	62.2	645	10	BG964218 602828996
7	246.4	62.2	1131	10	BG295496 602392826
8	244.8	61.8	560	10	BI794616 1c92e02.y
9	244.6	61.8	574	10	BI185580 UNT-P-FN-
10	234	59.1	501	9	AW211366 uo80b07.y
11	224.4	56.7	701	10	BI831943 603076873
12	223.4	56.7	336	9	AW326482 19106 MAR
13	222.8	56.3	817	10	BI1103329
14	195.4	49.3	546	10	BM488747 psm2n.pk0
15	191.6	48.4	453	10	BF916224 CM2-UT011
16	191.4	48.3	855	10	BI412175 602965377
17	188.4	47.6	967	10	BI739710 603361983

18	185.6	46.9	397	9	BE012485
19	182.2	46.0	697	10	BI854643
20	168	42.4	722	10	BI876051
21	166.2	42.0	622	10	BJ039504
22	166	42.0	681	9	AV704087
23	156.8	39.6	421	10	BI448355
24	149.8	37.8	637	9	AL654963
25	143.8	36.3	417	9	AA104365
26	142.6	36.0	546	9	BE014800
27	140.8	35.6	561	9	A1942584
28	139.4	35.2	678	9	BB654057
29	139.2	35.2	1295	11	AK015770
30	137.8	34.8	916	10	BI757180
31	129.6	32.7	567	10	BJ035232
32	129.6	32.7	612	10	BJ012200
33	128.2	32.4	512	9	AW210205
34	128	32.3	684	9	AI133696
35	128	32.3	750	10	BI831750
36	126.4	31.9	488	9	AI310713
37	126	31.8	309	9	AI630395
38	125.4	31.7	729	10	BI460337
39	116	29.3	803	10	BG777404
40	115.2	29.1	1209	11	AK006155
41	113.8	28.7	331	10	BI449976
42	109.8	27.7	666	9	BB644566
43	107.6	27.2	593	10	BG895490
44	104.8	26.5	483	9	AA061624
45	103.4	26.1	350	12	AZ693917

ALIGNMENTS

RESULT 1

BI827322

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI827322 747 bp mRNA linear EST 04-OCT-2001
603077989F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169694 5',
mRNA sequence.
BI827322
EST. GI:15938872
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 747)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11422 row: a column: 23
High quality sequence stop: 743.
Location/Qualifiers
1. .747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169694"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"

Site 1: Notif;
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT 188 a 206 c 201 g 151 t 1 others
ORIGIN

Query Match 70.7%; Score 280; DB 10; Length 747;
Best Local Similarity 100.0%; Pred. No. 6.2e-78;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgaccagctgccgcgtatcgcaatgtgttgagccctctgg 60
|||||
Db 84 ATGCAGAGCTCTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 143
Qy 61 aggattataagaacggaggccctatgaggccccctatgagggcgctgaacgctcacagcaaca 120
|||||
Db 144 AGGATTATAAGAACGAGGAGGCGCTATGGAGGCCCATGAGGGGCGCTGAACGTCACAGCAACA 203
Qy 121 ggcgcagggcctgccagcgcctttatttgcctgtctacgaaaaagttaaaaaagacattg 180
|||||
Db 204 GCGCAGGCGCTGCCAGCGCCCTTTATTTTGCCTCTACGAAAAAGTTAAAAAGACATTG 263
Qy 181 agtgatgtaataccaccctgggggcaatagccatattgccatattgctgcggcggtgtgtg 240
|||||
Db 264 ACTGATGTAATCCACCCTGGGGCAATAGCCATATTGCCAATGTTGCCATGGTGGCGCGGTGTGTG 323
Qy 241 gcaacattacttcattgatgcagccatgaacccctgcggaag 280
|||||
Db 324 GCAACATTACTTTCATGTCAGCCCATGAAACCCCTGCGGAAG 363

RESULT 2
AL530804
LOCUS AL530804 LTI_NFL001_NBC4 888 bp mRNA linear EST 13-FEB-2001
DEFINITION prime, mRNA sequence.
ACCESSION AL530804
VERSION AL530804.1 GI:12794297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 888)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. .888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD008VH17"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 168 a 244 c 322 g 139 t 15 others

ORIGIN

Query Match 70.1%; Score 277.6; DB 9; Length 888;
Best Local Similarity 99.3%; Pred. No. 3.8e-77;
Matches 278; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgaccagctgccgcgtatcgcaatgtgttgagccctctgg 60
|||||
Db 417 ATGCAGAGCTCTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 476
Qy 61 aggattataagaacggaggccctatgaggccccctatgagggcgctgaacgctcacagcaaca 120
|||||
Db 477 AGGATTATAAGAACGAGGAGGCGCTATGGAGGCCCATGAGGGGCGCTGAACGTCACAGCAACA 536
Qy 121 ggcgcagggcctgccagcgcctttatttgcctgtctacgaaaaagttaaaaaagacattg 180
|||||
Db 537 GCGCAGGCGCTGCCAGCGCCCTTTATTTTGCCTCTACGAAAAAGTTAAAAAGACATTG 596
Qy 181 agtgatgtaataccaccctgggggcaatagccatattgccatattgctgcggcggtgtgtg 240
|||||
Db 597 AGTATGTTATCCACCCTGGGGCAATAGCCATATTGCCAATGTTGCCATGGTGGCGCGGTGTGTG 656
Qy 241 gcaacattacttcattgatgcagccatgaacccctgcggaag 280
|||||
Db 657 GCAACATTACTTTCATGTCAGCCCATGAAACCCCTGCGGAAG 696

RESULT 3
BI045863
LOCUS BI045863 MR3-FN0209-300101-004-h10 FN0209 Homo sapiens cDNA, mRNA sequence.
DEFINITION BI045863
ACCESSION BI045863
VERSION BI045863.1 GI:14452485
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 480)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,F., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-MR3&t2-MR3-FN0209-300101-004-h10&t3=2001-01-30&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 480.

FEATURES
source
1. .480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0209"
/dev_stage="Adult"
/note="Organ: prostate_normal; Vector: puc18; Site_1: SmaI ; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
ORIGIN

116 a 126 c 146 g 92 t

Query Match 67.1%; Score 265.8; DB 10; Length 480;
Best Local Similarity 98.9%; Pred. No. 1.8e-73;
Matches 278; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 atgcagagctcacagcctgaccagctgcccgcctatcgcaatgtgttgagccctctgg 60
Db 74 ATGCAGAGCTTACAGCTCACCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 133
QY 61 aggattataagaacgagggcctatggagcccatggggctgaacgtcacagaaca 120
Db 134 AGGATTATTAAGAACGGAGGGCCTATGGAGGCCATGAGGGGCTGAACGTCACAGCAACA 193
QY 121 ggcgagggcctgcccacgcctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 194 GCGCAGGGCCCTGCCACCGCTTATTATTGCTGCTACGAAAGTTAAAAAGACATTG 253
QY 181 agtgatgtaaccacct-gggggcaatagccatattgccaatgtgctgcggcggtgtgt 239
Db 254 AGTGATGTAATCCACCTGGGGGCAATAGCCATATTGCCAATGTGTGCGCGGGTGTGT 313
QY 240 gcaacattactctatgatgcagccatgaacctgcggaag 280
Db 314 GGCACCATTTACTTCATGATGCAGCCATGAACCTTCGCGGAAG 354

RESULT 4
LOCUS B1541703 573 bp mRNA linear EST 30-AUG-2001
DEFINITION 455541 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION B1541703
VERSION B1541703.1 GI:15382815
KEYWORDS EST.
SOURCE cow.

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 573)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
G.D., Heaton,M.P., Laegreid,W., Rohrer,G.A., Chitko-McKown,C.G.,
Perteaux,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCAGCAG
Plate: 121 row: O column: 24
Seq primer: ATTAGGTGACACTATAG.

FEATURES
Source 1.573
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"

/clone_lib="MARC 1BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
BASE COUNT 121 a 166 c 176 g 110 t
ORIGIN

Query Match 64.6%; Score 256; DB 10; Length 573;
Best Local Similarity 94.6%; Pred. No. 2.4e-70;
Matches 265; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 atgcagagctcacagcctgaccagctgcccgcctatcgcaatgtgttgagccctctgg 60
Db 174 ATGCAGAGCTTACAGCCGATCCAGCCGCCCTATCGCAATGTGTGAGGCCCTCTGG 233
QY 61 aggattataagaacgagggcctatggagcccatggggctgaacgtcacagaaca 120
Db 234 AGGATTATTAAGAACGGAGGGCCTGTGGAGGCCATGCGAGGCTGAACGTCACAGCAACA 293
QY 121 ggcgagggcctgcccacgcctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 294 GCGCAGGGCCCTGCCACCGCTTATTATTGCTGCTATGAAAAGTTAAAAAGACATTG 353
QY 181 agtgatgtaaccacctgggggcaatagccatattgccaatgtgctgcggcggtgtgtg 240
Db 354 AGTGATGTAATCCACCTGGGGGCAATAGCCATATTGCCAATGTGTGCGCGGGTGTGTG 413
QY 241 gcaacattactctatgatgcagccatgaacctgcggaag 280
Db 414 GCAACATTACTTCATGATGCTGCCATGAATCCAGTGAAG 453

RESULT 5
LOCUS BE913718 598 bp mRNA linear EST 29-SEP-2000
DEFINITION 60166072F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5',
mRNA sequence.

ACCESSION BE913718
VERSION BE913718.1 GI:104111618
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 598)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM9146 row: o column: 17
High quality sequence stop: 593.

FEATURES
Source 1.598
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969256"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

BASE COUNT	147 a	189 c	193 g	116 t


```

Best Local Similarity 92.5%; Pred. No. 3.5e-67;
Matches 259; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY .1 atcgaggtctacagctgaccagctgccgctatcgcaatgtgttgagccctctgg 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 261 ATGCAGAGCTACAGCTACCCAGCGCCGCTATCGGAACGTGTGGAGGCTCTCTGG 320
QY 61 aggattataagaacgagggccttatggagcccatgagggggtggaactcacagcaaca 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 321 AGAATCATGAGGACAGAGGCGCTGTGGAGGCCATCGGGGGCTGAACGTCACAGCAACA 380
QY 121 ggcgagggcctgcccagccctttatttgcctgctagcaaaagttaaaagacattg 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 381 GCGCGGGGCGCTGCCACCGCCCTCTATTGTGCTGTACGAAAGTAAAAAGACATTG 440
QY 181 agtgatgtaatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtg 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 441 AGTGAGGTAATCCACCCAGCGGCGCATATGTCATTTGCCAATGGTCAGCGCGATGTGTG 500
QY 241 gcaacattacttcgatgcagccatgaacccctgcggaag 280
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 501 GCGACATTACTTCATGATGAGCCATGAATCCAGCGGAAG 540

RESULT 8
BI794616 560 bp mRNA linear EST 01-OCT-2001
LOCUS ic92e02.y1 Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus
DEFINITION musculus DNA 5' similar to TR:Q23125 Q23125 W02B12.9 PROTEIN. [1]
; mRNA sequence.
ACCESSION BI794616
VERSION BI794616.1 GI:15822341
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 560)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., Cole,R., Cole,R., Tsagarisvill,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: ic92e02.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 430.
Location/Qualifiers
1..560
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1
NI-MMS1"
/sex="Both for embryonic & newborn, male for adult and
adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
adult, mixed"
/lab_host="DH10B"

```

/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; Five libraries representing E10.5/12.5 pancreatic bud, E16.5 pancreas, newborn pancreas, adult pancreas, and adult islets of Langerhans were separately constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA was made by oligo-dT priming and size-selected by column fractionation. Libraries were amplified once on solid support and plasmid DNA from each library was prepared and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an EcoT of 6. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 134 a 158 c 163 g 105 t
ORIGIN

Query Match 61.8%; Score 244.8; DB 10; Length 560;
Best Local Similarity 92.1%; Pred. No. 8.8e-67;
Matches 258; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 atcgaggtctacagctgaccagctgccgctatcgcaatgtgttgagccctctgg 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84 ATGCAGAGCTACAGCTACCCAGCGCCGCTATCGGAACGTGTGGAGGCTCTCTGG 143
QY 61 aggattataagaacgagggccttatggagcccatgagggggtggaactcacagcaaca 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 144 AGAATCATGAGGACAGAGGCGCTGTGGAGGCCATCGGGGGCTGAACGTCACAGTAACA 203
QY 121 ggcgagggcctgcccagccctttatttgcctgctagcaaaagttaaaagacattg 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 204 GCGCGGGGCGCTGCCACCGCCCTCTATTGTGCTGTACGAAAGTAAAAAGACATTG 263
QY 181 agtgatgtaatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtg 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 264 AGTGAGCTAATCCACCCAGCGGCAATGCCATATTGCCAATGGTCAGCGCGATGTGTG 323
QY 241 gcaacattacttcgatgcagccatgaacccctgcggaag 280
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 324 GCGACATTACTTCATGATGAGCCATGAATCCAGCGGAAG 363

RESULT 9
BI185580
LOCUS BI185580 574 bp mRNA linear EST 10-JUL-2001
DEFINITION UNL-P-FN-cy-f-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
UNL-P-FN-cy-f-02-0-UNL 3', mRNA sequence.
ACCESSION BI185580
VERSION BI185580
KEYWORDS EST. GI:14659989
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 574)
Caetano,A.R., Johnson,R.K. and Pomp,D.
Generation and sequence characterization of a normalized cDNA
library from swine ovarian follicles
Unpublished (2001)
JOURNAL Contact: Pomp, D
COMMENT Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message.
Seq primer: M13 -29
POLYA=No.

cdna Library Preparation: Life Technologies, Inc.
 cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM11419 row: a column: 20
 High quality sequence start: 29
 High quality sequence stop: 699.
 Location/Qualifiers
 FEATURES
 source
 1..701
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5168539"
 /clone_lib="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."
 BASE COUNT 176 a 167 c 208 g 150 t

Query Match 56.7%; Score 224.4; DB 10; Length 701;
 Best Local Similarity 99.6%; Pred. No. 2.9e-60;
 Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 atgcagagctcacagcctgacccagctccgctatcgcaatgtgttgaggccctctgg 60
 Db 302 ATGCAGAGCTTACAGCTGACCCAGCTCCGCTATCCATGTTGTGAGGCCCTCTGG 361
 Qy 61 aggtattataagaagcggagggcctatggaggcccatggggcctgaaagcgtcacagcaaca 120
 Db 362 AGGATTATAAGAAGAGGAGGCGCTATGAGGCGCCATGAGGGCTGAACGTACAGCAACA 421
 Qy 121 ggcagggcctgccagcgcctttatttgcctgctacgaaagttaaaaagacattg 180
 Db 422 GCGCAGGGCCTGCCAGCGCCTTATTTCCTGCTACGAAAAGTTAAAAAGACATTG 481
 Qy 181 agtgaatgaatccaccctgggggcaatagccatatgtgccaatggtg 226
 Db 482 AGTGATGTAATCCACCCCTGGGGGCANTAGCCATATTGCCAATGCTG 527

RESULT 12
 AW326482
 LOCUS 336 bp mRNA linear EST 25-APR-2001
 DEFINITION 19106 MARC 1BOV Bos taurus cdna 5', mRNA sequence.
 ACCESSION AW326482
 VERSION AW326482.1 GI:6762403
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 336)
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,T.E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
 ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G.,
 Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
 Keele,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cdna
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013

COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACAGTCACGACG
 Plate: 8 row: F column: 16
 Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

1..336
 /location/Qualifiers
 1..336
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 1BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."
 BASE COUNT 72 a 96 c 107 g 61 t

ORIGIN

Query Match 56.4%; Score 223.4; DB 9; Length 336;
 Best Local Similarity 95.4%; Pred. No. 4.7e-60;
 Matches 230; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 Qy 1 atgcagagctcacagcctgacccagctccgctatcgcaatgtgttgaggccctctgg 60
 Db 96 ATGCAGAGCCTTACAGCGGATCCAGCCGCCGCTATCGCAATGTTGTGAGGCCCTCTGG 155
 Qy 61 aggtattataagaagcggagggcctatggaggcccatggggcctgaaagcgtcacagcaaca 120
 Db 156 AGGATTATAAGAAGAGGAGGCGCTGTGAGGCGCCATGCGAGGCTGAGGCTCAGCAACA 215
 Qy 121 ggcagggcctgccagcgcctttatttgcctgctacgaaagttaaaaagacattg 180
 Db 216 GCGCAGGGCCTGCCAGCGCCTTATTTCCTGCTATGAAAAGTTAAAAAGACATTG 275
 Qy 181 agtgaatgaatccaccctgggggcaatagccatatgtgccaatggtgcccgggtgtgtg 240
 Db 276 AGTGATGTAATCCACCCCTGGGGGCANTAGCCATATTGCCAATGTTGGCTGGTGTGTG 335
 Qy 241 g 241
 Db 336 G 336

RESULT 13

BI103329

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI103329 817 bp mRNA linear EST 26-JUN-2001
 60288838F1 NCI_CGAP_Kid14 Mus musculus cdna clone IMAGE:5043992
 5', mRNA sequence.
 BI103329
 BI103329.1 GI:14554222
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 817)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health.
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cdna Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1120 row: 1 column: 09
 High quality sequence stop: 777.

FEATURES
 source
 1..817
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5043992"
 /clone_lib="NCI_CGAP_Kid14"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: Kidney; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library. 1"
 BASE COUNT 147 a 222 c 316 g 132 t
 ORIGIN
 Query Match 56.3%; Score 222.8; DB 10; Length 817;
 Best Local Similarity 91.5%; Pred. No. 9.9e-60;
 Matches 258; Conservative 0; Mismatches 22; Indels 2; Gaps 2;

QY 1 atcgaggtctacagctgacccagctgcccgtatgcagctatgcagctgtgtgagccctctgg 60
 Db 422 ATGCAGAGCTACAGCTGACCCAGCCGCCGCTATCGGAACGTGTGGAGGCTCTCTGG 481
 QY 61 aggattataagaacggaggccctatgagggccatgaggggctgaaactcacagcaaca 120
 Db 482 AGAATCATGAGACAGAGGGGCTCTGGAGGCCCATGCGGGGCTGAACGTACACAGCAACA 541
 QY 121 ggcgagggcctgccacgccccttatttgcctgctacgaaagttaaaaaagacattg 180
 Db 542 GGCAGGGGCTGCGCCAGCCGCTCTATTTGCTCTACGAAAAGTTAAAAAGACATTG 601
 QY 181 agtataatccaccctgggggcaatagccattatcccaattg-tgcggccgggtgtgt 239
 Db 602 AGTACGTAATCCACCCAGGGGGCAATAGCCATATTCCTATGTTGACGCGGATGTGT 661
 QY 240 ggcacattac-ttcattgatgcagccatgaacctgcggaag 280
 Db 662 GGCACATTACTTTTCATGATGCAGGCATGAATCCAGCGGAAG 703

RESULT 14
 BM488747
 LOCUS
 DEFINITION
 pgm2n.pk008.12 Normalized Chicken Breast Muscle, Leg Muscle, and
 Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
 clone pgm2n.pk008.12 5' similar to ref:XP_050766.1 (XM_050766)
 putative mitochondrial solute carrier [Homo sapiens]
 gb|AAK49519.1|AF327402.1 (AF327402) putative mitochondrial solute
 carrier splice variant [Homo sapiens], mRNA sequence.
 ACCESSION
 BM488747
 VERSION
 BM488747.1 GI:18609678
 KEYWORDS
 EST.
 SOURCE
 chicken.
 ORGANISM
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE
 1 (bases 1 to 546)
 Cogburn,L.A. and Monsonego-Ornan,E.
 ESTs from Normalized Chicken Breast
 Epiphyseal Growth Plate cDNA library, USDA/IRAFS Animal Genome
 Project.
 JOURNAL
 Unpublished (2002)
 COMMENT
 Contact: Larry A. Cogburn
 University of Delaware

Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2822
 Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES
 source
 1..546
 Location/Qualifiers

/organism="Gallus gallus"
 /strain="Commercial broiler and Ottawa Res. Centre
 Strains 90 & 21"
 /db_xref="taxon:9031"
 /clone="pgm2n.pk008.12"
 /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
 and Epiphyseal Growth Plate cDNA library (pgm2n)"
 /sex="Male and Female"
 /tissue_type="Breast muscle, leg muscle and epiphyseal
 growth plate"
 /dev_stage="Breast,leg;Embryo(d19);post-hatch(1d,1,3,5,7,9
 ,11 weeks);growth plate(1d,7d,14d post-hatch)"
 /lab_host="E. coli EMD108"
 /note="Vector: pCMVSPORT6; Library made from equivalent
 pools of total RNA isolated from each tissue (embryonic
 muscle 33.3%, juvenile muscle 33.3%, and epiphyseal growth
 plate 33.3% of the final RNA pool). Single pass sequencing
 from 5'-end"
 BASE COUNT 90 a 183 c 197 g 70 t 6 others
 ORIGIN

Query Match 49.3%; Score 195.4; DB 10; Length 546;
 Best Local Similarity 80.7%; Pred. No. 4.5e-51;
 Matches 226; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 1 atcgaggtctacagctgacccagctgcccgtatgcagctgtgtgagccctctgg 60
 Db 217 ATGCAGAGCTGCGGCCCGAGCCGCCGCTACCGAACGTGTGGAGGCCCTGTGG 276
 QY 61 aggattataagaacggaggccctatgagggccatgaggggctgaaactcacagcaaca 120
 Db 277 CGCATCTGCTACCGAGGGCGTGTGGAGGCCCATGCGGGGCTGAACATCACGCCACC 336
 QY 121 ggcgagggcctgccacgccccttatttgcctgctacgaaagttaaaaaagacattg 180
 Db 337 GGCAGGGGCTGCGCCAGCCGCTCTACTTGCCTGTACGAAAAGTTAAAAAGACGCTG 396
 QY 181 agtataatccaccctgggggcaatagccattatcccaattg-tgcggccgggtgtgt 240
 Db 397 AGCAGCTCATCCACGGGGGGCAATAGCCATGTGCGNNCGTGCGAGCGGGGTGTGTA 456
 QY 241 gcaacattacttcattgatgcagccatgaacctgcggaag 280
 Db 457 GCAACATTGCTCCACGACGCGCAGCGATGAACCCCTGCAGAAG 496

RESULT 15

BF916224
 LOCUS
 DEFINITION
 CM2-UT0112-091200-598-c06 UT0112 Homo sapiens cDNA, mRNA sequence.
 ACCESSION
 BF916224
 VERSION
 BF916224.1 GI:12307682
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 453)
 AUTHORS
 Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 TITLE
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

Job time: 0191 sec

THIS PAGE BLANK (USPTO)